昭和38年7月2日第3種郵便物認可
2010年4月1日発行　第48巻4号　毎月1回1日発行
ISSN 0386—2240

数理科学

4
2010

特集　現代数学はいかに使われているか［確率編］



No.562

サイエンス社

MATHEMATICAL SCIENCES
Number 562 April 2010

現代数学はいかに使われているか［確率編］

表紙：戸川　隼人

お得な定期購読をおすすめします

毎号確実に御手元へ

地域によって書店店頭に陳列されない場合があります．**郵送料小社負担**にて確実に御手元へ届く定期予約購読をおすすめします．

予約購読は1年間（本誌のみ）11,000円，1年間（本誌＋別冊2冊）14,500円，半年間（本誌のみ）5,500円，半年間（本誌＋別冊1冊）7,250円です．（税込み価格）

造本コストは更に上昇気味です．小社としてもあらゆる努力をもって廉価で提供してゆきたいと思っておりますが，定価改定の止むなき事態がかんがえられます．この場合予約購読期間中のものにつきましては料金据置きとなります．なお，臨時別冊のご注文については，本誌予約購読者の方には，商品到着後のお支払いにて承ります．

お申し込みは本誌とじ込みの振替用紙を御利用下さい．

数理科学特集一覧

63年/7～2007年/5 省略

2007年/6 物理における位相の世界
/7 代数的物理観
/8 数学における場の量子論
/9 ディラック
/10 フーリエ解析の力
別冊 素粒子物理学の発展
/11〈電弱統一理論〉
/12「時空」への問いかけ
2008年/1 現象を視る
/2 ゲージ重力対応
/3 現代物理のための数学キーワード
/4 現代数学はいかに使われているか［解析編］
別冊 多様体の広がり
/5 物理と数学の難所
/6 線形代数の力
/7 物理的思考法のすすめ
/8 数の魅力
/9 トポロジカルな弦の世界
/10 ポアンカレ
別冊 対称性と保存則
/11 物理と論理
/12 確率論的物理観
2009年/1 グラフィックでわかる物理
/2 発展する量子力学
/3 方程式に潜む対称性
/4 現代数学はいかに使われているか［幾何編］
別冊 代数学の魅力
/5 微積分を楽しむ
/6 現象と物理概念を結ぶ方程式
/7 現代物理の世界像
/8 複素数
/9 シュレーディンガー
/10 数学と物理に広がる不変量
別冊 量子重力理論
/11 可積分系の世界
/12《原理》の探求
2010年/1 無限次元の魅力
/2 数学の可視化
/3 物理数学の効用

数理科学臨時別冊一覧

08年/7 ネットワーク科学への招待
08年/9 物理のためのリー群とリー代数
08年/11 量子暗号理論の展開
09年/1 量子力学の考え方
09年/7 ゲージ理論の発展
09年/9 量子コンピュータの基礎［第2版］
09年/11 重点解説基礎微分幾何
09年/12 ε–δ 論法再入門
10年/1 多彩な量子の世界
10年/3 超関数・フーリエ変換入門

「数理科学」のバックナンバーは下記の書店・生協の自然科学書売場で特別販売しております

紀伊國屋書店本店（新宿）
紀伊國屋書店新宿南店
あおい書店中野本店
オリオン書房ノルテ店（立川）
くまざわ書店本店（八王子）
くまざわ書店桜ヶ丘店（多摩）
書泉グランデ（神田）
書泉ブックタワー（秋葉原）
三省堂本店（神田）
リブロ池袋店
ジュンク堂書店（池袋）
ジュンク堂書店（新宿）
芳林堂（高田馬場）
八重洲ブックセンター（東京駅前）
丸善丸の内本店（東京駅前）
ブックファースト青葉台店（横浜）
栄松堂ジョイナス店（横浜）
有隣堂イセザキ町店（横浜）
有隣堂西口・東口（横浜）
有隣堂厚木店（厚木）
文教堂ブックセンター（川崎）
ジュンク堂書店（盛岡）
ジュンク堂書店（新潟）
旭屋船橋店
戸田書店静岡本店
ジュンク堂大阪本店
紀伊國屋書店梅田店（大阪）
旭屋本店（大阪）
アバンティブックセンター（京都）
ジュンク堂書店（京都）
ジュンク堂書店（神戸）
ジュンク堂三宮駅前店（神戸）
ジュンク堂書店（大分）
喜久屋書店倉敷店
紀伊國屋書店福岡店
紀伊國屋書店福岡本店
ジュンク堂福岡店
紀伊國屋書店新潟店
金港堂（仙台）
金港堂パーク店（仙台）
ジュンク堂書店（仙台）
鹿島ブックセンター（いわき）
友朋堂（つくば）

――大学生協・売店――
東京大学 本郷・駒場
東京工業大学 大岡山・長津田
東京理科大学 新宿
早稲田大学 理工学部
慶応義塾大学 矢上台
福井大学
筑波大学 大学会館書籍部 第一学群書籍部
埼玉大学
名古屋工業大学・愛知教育大学
大阪大学・神戸大学 ランス
京都大学・九州工業大学
東北大学 理薬・工学
室蘭工業大学
徳島大学 常三島
愛媛大学 城北
山形大学 小白川
島根大学
学習院大学 成文堂
北海道大学 クラーク店

April 2010 MATHEMATICAL SCIENCES Number 562

特集／現代数学はいかに使われているか［確率編］

巻頭言

楠岡 成雄

皆既日食をいつどこで見ることができるかは，100年先であっても正確に予測できるようである．しかし，来年のプロ野球の結果を正確に予測することは不可能であろう．正確に予測することができない現象（不確実な現象）を解析するときには，多くの場合に確率の概念が用いられる．「確率とは何か」ということは現在でも哲学の研究対象であり，簡単な問題ではない．しかし数学では20世紀前半にこの哲学的な問題を棚上げにし，どのような前提からどのような結論が数学的に導かれるか，ということのみを研究の対象としてきた．その後約80年間に数学としての確率論は飛躍的に発展した．特に20世紀後半における確率解析の発展は著しいものであった．発展の過程で確率論は諸科学からの影響を受け，純粋な数学的直観に基づいた確率モデル・確率過程モデルのみならず，応用を意識した様々なモデルが見出された．そして諸科学にその成果を還元・提供できるようになった．

今日では，自然科学，社会科学を問わず多くの分野で確率モデルが用いられている．確率論の基礎が構築された際に，物理学，特に統計力学が強く意識された．統計力学は確率モデルが最も盛んに使われている分野であり，確率論に対して常に研究対象や問題を提供している分野である．生物学では，遺伝におけるメンデルの法則を初め，確率モデルが古くから用いられている．19世紀終わりから数理統計学が急速に発展したが，その研究の源は農学・医学を初めとする生命科学であった．情報科学でも確率モデルが用いられるが，特にシャノンのエントロピーは，数学に対して大きな影響を及ぼした．その他の自然科学や工学においても確率モデルがよく用いられている．また，経済学では，経済主体の意思決定と関係して「確率」が意識されてきた．今日では極めて複雑な確率過程モデルがしばしば用いられている．特に，ファイナンスでは確率過程モデルはその理論の基礎そのものを与えている．

この特集では様々な分野の研究者に確率論・確率モデルがどのように用いられているかについて執筆していただいた．確率論の応用の様子を知ることができたと読者の方々に感じていただければ幸いである．筆者は確率論研究者であるが，応用についてはごく一部のことしか知らない．筆者自身も記事を読むことを楽しみにしている．確率モデルが用いられる代表的な分野を選んだつもりであるが，偏りがあるとすれば，筆者の無知によるものである．このことは，あらかじめお断りしておく．

確率モデルを用いて現実の現象を解析するときには，「確率とは何か」という問題が再び現れる．確率論の最も重要な研究成果として，大数の法則，中心極限定理などの原理がある（数学ではしばし

ば極限定理と呼ばれている）．これら極限定理は当初，独立変数の和などの限られたモデルに対してのみ定理として証明されたが，今日では多くの構成要素があり，それらが高い独立性を持つような様々なモデルに対して定理として証明されている．これらの古典的原理が成立する場合は，集計効果により，個々の要素の特性が全体に影響を及ぼすことがなくなり，その結果「確率とは何か」をあまり深刻に考える必要がなくなる．物理学などでは，これらの原理が多くの場合に成立すると信じられている．

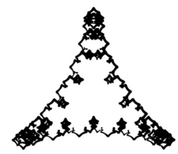
一方，構成要素の数が十分多くない場合や，多くても独立性が高くない場合は，簡単な形では極限定理が成立せず，構成要素の個々の特性が全体に影響を及ぼすことも数学的に示されている．そのような場合，確率モデルを適用する際には，そもそも，そこでいう確率とは何かを真剣に考える必要がある．確率モデルがパラメータを含む場合はそのパラメータを決めるのに統計的推測が必要となる．その確率モデルがしっかりとした根拠を持つものであれば，過去のデータに基づきパラメータを推定することは可能である．しかしながら今日では，複雑かつ不確実な現象を解析するのに，とりあえず確率モデルを立てて考えるということも少なくない．その場合には，安易にそのモデルに確率論や統計学の理論を適用し，その結論が数学的な理論に保証された堅固なものだと考えるならば，大きな失敗を招く．

確率モデルを適用する範囲が広がりつつある現在，用いているモデルが現実とは大きく違っている可能性を考える必要性も現れつつある．それを扱うのは，「リスク管理」と呼ばれる分野であろうが，理論的には未成熟な発展途上の分野であり，確率論・統計学との整合性も極めて曖昧である．20世紀初頭の「確率論」と同じく，数学としての「リスク管理理論」を作ることが21世紀の課題であると感じている．しかし，20世紀の「確率論」のような大成功が収められるであろうか．

（くすおか・しげお，東京大学大学院数理科学研究科）

統計力学と確率論

調和振動子（自由場の理論）からランダムウォークへ



田崎　晴明

1. 統計力学と確率

ミクロな（小さな）原子や分子の性質をもとに，それらが数多く集まってできているマクロな（大きな）物質の性質を調べるのが統計力学である．昨年出版したばかりの私の教科書[1]に

> 簡単化し過ぎることを恐れずに言い切れば，**マクロな熱力学の体系と整合するように，ミクロな（量子）力学の体系に確率分布を導入したのが，平衡統計力学**なのだ（p.81）．

と書いたように，統計力学というのは確率論に基づいた体系だ．だから確率論が活躍する場面は無数にあるのだが，ここでは，「バネがたくさん集まった系の統計力学」とランダムウォークの関わりについてじっくりと解説しよう．多体系の統計力学や場の量子論の問題を，確率幾何的な問題に置き換えて解析するという，現代の物理学にくり返し登場する手法のひな形でもある．

といっても，統計力学の知識や高度な確率論は仮定しない．高校生にもわかる部分積分だけを武器に，2 つの分野の関わりを描いてみたい．

2. 一つの調和振動子の統計力学

簡単な場合からはじめよう．統計力学をご存知ない読者には即席の「統計力学入門」にもなる．

受験問題に出てくるような，理想的なバネの先端に小さな球がくっついた系を考える．この系を大学風に調和振動子と呼ぼう．バネの自然の状態からの「伸び」を変数 $\varphi \in \mathbb{R}$ で表す[*1)]．φ が正のときはバネは自然状態より伸びており，φ が負のときは自然状態より縮んでいる．

伸びが φ のときにバネに蓄えられるポテンシャルエネルギーを

$$V(\varphi) = \frac{\kappa}{2}\varphi^2 \tag{1}$$

とする．$\kappa > 0$ はバネ定数である．(1) は，伸びが φ のときに小さな球に働くバネの復元力が $-\kappa\varphi$ であることを表している．

この調和振動子が絶対温度 T の環境に置かれているとしよう．例えば，一定温度の水中に小さなバネに結ばれた小さな球が置かれていることを想像すればいい[*2)]．このようなとき，小球は（例えば周囲の水分子が次々と衝突することで）不規則な力を受けてランダムな運動をする．だが，このランダムな運動には統計的な規則性がある．ある時刻に小球がどの位置にあるかを正確にいうことはできないが，十分に時間が経った**平衡状態**では，バネの伸びが φ である確率（密度）が (1) のポテンシャルを使って $p(\varphi) \propto \exp[-V(\varphi)/(k_{\mathrm{B}}T)] = \exp[-(\mu/2)\varphi^2]$

*1)　本来，φ は正にも負にも有限の範囲しか動かないはずだが，簡単のため φ は実数全体を動くとする．これから考察する平衡状態では，$|\varphi|$ がある程度以上大きくなる確率は極めて小さいので，この理想化はさしたる問題ではない．

*2)　水中で，直径 1 ミクロン程度のプラスティック球をレーザーで作った「バネ」でつかまえるといった実験は今日では簡単に行なえる．

と表されることが知られている．これが平衡統計力学の基礎となる原理である．ボルツマン定数 $k_{\mathrm{B}} \simeq 1.38 \times 10^{-23}$ J/K は絶対温度をエネルギーに換算するための基本的な定数である．以下では温度 T は一定とし，$\mu := \kappa/(k_{\mathrm{B}}T)$ を系を特徴付けるパラメターとしよう．

$f(\varphi)$ を φ の任意の実関数とする．上に述べた平衡状態でのこの関数の期待値（平均値）は，確率密度で重みをつけて積分することで，

$$\langle f(\varphi)\rangle = \frac{\displaystyle\int_{-\infty}^{\infty} d\varphi\, f(\varphi)\, \exp[-(\mu/2)\varphi^2]}{\displaystyle\int_{-\infty}^{\infty} d\varphi\, \exp[-(\mu/2)\varphi^2]} \quad (2)$$

となる．様々な物理量（関数）の平衡状態での期待値を求めることが，統計力学の応用での主要な仕事である．

例題（および，この先の問題の準備）として $f(\varphi) = \varphi^2$ の場合を考えよう*3)．(2) の分子の積分を書き下し，部分積分を使って変形すると，$\int_{-\infty}^{\infty} d\varphi\, \varphi^2 \exp[-(\mu/2)\varphi^2] = -\mu^{-1}\int_{-\infty}^{\infty} d\varphi\, \varphi\,(\exp[-(\mu/2)\varphi^2])' = \mu^{-1}\int_{-\infty}^{\infty} d\varphi\, \exp[-(\mu/2)\varphi^2]$ となる．最後に現れた積分は (2) の分母そのものだから，結局 $\langle\varphi^2\rangle = \mu^{-1}$ が得られた．

3. 格子上の調和振動子の系

この解説の本題である格子上の調和振動子の系を定義しよう．簡単に言えば，立方格子の各々の格子点に，前節で考えたようなバネ＋小球の系が「乗っている」状況を扱う．

まず舞台となる格子から（図 1）．Λ を一辺が L の d 次元立方格子の格子点の集合とする．格子点とは d 次元ユークリッド空間の点 $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ であり，座標成分は $x_j \in \{0, \pm1, \pm2 \ldots \pm (L-1)/2\}$ の範囲を動く．また $\mathcal{B}$ を互いに隣り合った（つまり，距離が 1 の）格子点の組 $\{\boldsymbol{x}, \boldsymbol{y}\}$ すべての集合とする（$\boldsymbol{x}$, $\boldsymbol{y}$ の順番は気にしない）．周期的境界条件を課し，格子のちょうど反対側の端に位置する 2 つの格子点も「隣り合っている」とみなす（細かいことなので初学者は気にしなくてよい）．

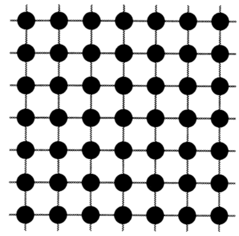


図 1　$d = 2, L = 7$ の場合の格子．Λ の要素の格子点が黒丸であり，$\mathcal{B}$ の要素の格子点の組を線で表した．

各々の格子点 $\boldsymbol{x} \in \Lambda$ に調和振動子の系を対応させる．格子点 $\boldsymbol{x}$ に対応するバネの「伸び」を $\varphi_{\boldsymbol{x}}$ としよう．こうすると，全部で L^d 個の実変数を扱うことになる．これらをすべてまとめたものを**配位**と呼び $\Phi = (\varphi_{\boldsymbol{x}})_{\boldsymbol{x}\in\Lambda} \in \mathbb{R}^{L^d}$ と書く．Φ はすべてのバネの伸びを表している．

配位 Φ が与えられたときの系全体のポテンシャルエネルギーを

$$\mathcal{V}[\Phi] := \sum_{x\in\Lambda} \frac{\kappa}{2}(\varphi_{\boldsymbol{x}})^2 + \sum_{\{x,y\}\in\mathcal{B}} \frac{\kappa'}{2}(\varphi_{\boldsymbol{x}} - \varphi_{\boldsymbol{y}})^2 \quad (3)$$

としよう．第一項は前節のポテンシャル (1) を全格子点について足しあげたもの．これだけでは面白くないので，第二項では，隣り合った格子点での $\varphi_{\boldsymbol{x}}$ の値を近づける傾向をもった，もう一種類の「バネ」の効果を取り入れた．こちらのバネ定数を $\kappa' > 0$ とした．

(3) のようなポテンシャルエネルギーを実現するようにバネや小球を配置するのは必ずしも容易ではない．これは，現実的な物理系のモデルというよりも，調和型のポテンシャルで結ばれた大自由度系の性質を知るための理想的なモデルとみなしたほうがいいだろう．実際，このような系は，バネとは無縁のスピン系や場の量子論にも顔を出す．

さて，この系が絶対温度 T の環境に置かれているとしよう．やはり配位 Φ はランダムな変動を見せるが，平衡状態では，配位 Φ が現れる確率密度は $\exp[-\mathcal{V}[\Phi]/(k_{\mathrm{B}}T)]$ に比例する．後の便利のた

*3)　より簡単な例題として，$\langle\varphi\rangle = 0$ となることは $\varphi \exp[-(\mu/2)\varphi^2]$ が奇関数であることからすぐにわかる．

め，$\mathcal{S}[\Phi] := \mathcal{V}[\Phi]/(k_{\mathrm{B}}T)$ とする．(3) の第二項の積を展開し，対角項を第一項とまとめれば，

$$\mathcal{S}[\Phi] := \sum_{\boldsymbol{x}\in\Lambda}\Big(\frac{\mu}{2}+d\,\nu\Big)(\varphi_{\boldsymbol{x}})^2 - \sum_{\{\boldsymbol{x},\boldsymbol{y}\}\in\mathcal{B}} \nu\,\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}} \tag{4}$$

となる．ここで，$\nu := \kappa'/(k_{\mathrm{B}}T)$ とした．

$F[\Phi]$ を，配位 $\Phi = (\varphi_{\boldsymbol{x}})_{\boldsymbol{x}\in\Lambda}$ の任意の実数値関数とする．平衡状態での $F[\Phi]$ の期待値を，(2) と全く同様に

$$\langle F[\Phi]\rangle := \frac{\displaystyle\int \mathcal{D}\Phi\, F[\Phi]\, e^{-\mathcal{S}[\Phi]}}{\displaystyle\int \mathcal{D}\Phi\, e^{-\mathcal{S}[\Phi]}} \tag{5}$$

と定義する．$e^{-\mathcal{S}[\Phi]}$ という重みをかけてすべての可能な配位について積分するだけだが，L^d 個の $\varphi_{\boldsymbol{x}}$ すべてについて積分することに注意．ここで

$$\int \mathcal{D}\Phi(\cdots) := \Big(\prod_{\boldsymbol{x}\in\Lambda}\int_{-\infty}^{\infty} d\varphi_{\boldsymbol{x}}\Big)(\cdots) \tag{6}$$

は多重（L^d 重）積分の略記である．

4. 部分積分から漸化式へ

2つの任意の格子点 $\boldsymbol{x},\boldsymbol{y}\in\Lambda$ について，期待値 $\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle$ を考えよう．この期待値のことを**二点相関関数**と呼ぶ．そもそも積分の範囲が無限なので，積分が収束するかどうかも自明ではないが，任意の $\mu>0, \nu>0$ について $|\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle| < \infty$ は簡単に証明できるので，それは認めて話を進めよう．

一変数のときと同様の部分積分をこの多変数積分についても行なう．まず，偏微分の計算をする．

$$\frac{\partial}{\partial\varphi_{\boldsymbol{x}}}\mathcal{S}[\Phi] = (\mu+2d\,\nu)\varphi_{\boldsymbol{x}} - \nu\sum_{\boldsymbol{z}\in\mathcal{N}(\boldsymbol{x})}\varphi_{\boldsymbol{z}} \tag{7}$$

ここで，$\mathcal{N}(\boldsymbol{x}) := \{\boldsymbol{z}\in\Lambda\,|\,\{\boldsymbol{x},\boldsymbol{z}\}\in\mathcal{B}\}$ は格子点 $\boldsymbol{x}$ と隣り合う格子点の集合である．これを用いて，部分積分すれば

$$\int\mathcal{D}\Phi\left\{(\mu+2d\,\nu)\varphi_{\boldsymbol{x}} - \nu\sum_{\boldsymbol{z}\in\mathcal{N}(\boldsymbol{x})}\varphi_{\boldsymbol{z}}\right\}\varphi_{\boldsymbol{y}}\,e^{-\mathcal{S}[\Phi]}$$
$$= -\int\mathcal{D}\Phi\,\varphi_{\boldsymbol{y}}\frac{\partial}{\partial\varphi_{\boldsymbol{x}}}e^{-\mathcal{S}[\Phi]} = \delta_{\boldsymbol{x},\boldsymbol{y}}\int\mathcal{D}\Phi\,e^{-\mathcal{S}[\Phi]} \tag{8}$$

となる[*4]．$\partial\varphi_{\boldsymbol{y}}/\partial\varphi_{\boldsymbol{x}} = \delta_{\boldsymbol{x},\boldsymbol{y}}$ を用いた．最右辺の積分は (5) の分母そのものなので，期待値についての関係

$$(\mu+2d\,\nu)\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle - \nu\sum_{\boldsymbol{z}\in\mathcal{N}(\boldsymbol{x})}\langle\varphi_{\boldsymbol{y}}\varphi_{\boldsymbol{z}}\rangle = \delta_{\boldsymbol{x},\boldsymbol{z}} \tag{9}$$

が出る．整理すると，二点相関関数が満たす漸化式

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\,\delta_{\boldsymbol{x},\boldsymbol{y}} + \frac{\zeta}{2d}\sum_{\boldsymbol{z}\in\mathcal{N}(\boldsymbol{x})}\langle\varphi_{\boldsymbol{z}}\varphi_{\boldsymbol{y}}\rangle \tag{10}$$

が得られる．ただし，$g_0 = (\mu+2d\nu)^{-1}$ および $\zeta = 2d\nu/(\mu+2d\nu) < 1$ とした．

5. 漸化式からランダムウォークへ

漸化式 (10) の意味を探るため，$\boldsymbol{x}$ と $\boldsymbol{y}$ が図 2 (a) のように距離 2 だけ離れているとしよう．このときは明らかに $\boldsymbol{x}\neq\boldsymbol{y}$ なので $\delta_{\boldsymbol{x},\boldsymbol{y}} = 0$ であり，(10) は

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = \frac{\zeta}{2d}\sum_{\boldsymbol{x}_1\in\mathcal{N}(\boldsymbol{x})}\langle\varphi_{\boldsymbol{x}_1}\varphi_{\boldsymbol{y}}\rangle \tag{11}$$

となる．右辺での和は $\boldsymbol{x}$ と隣り合う格子点すべてについてとる（図 2 (a)）．後の便利のため，足しあげる格子点を $\boldsymbol{z}$ でなく $\boldsymbol{x}_1$ と呼んだ．

次に，(11) 右辺の $\langle\varphi_{\boldsymbol{x}_1}\varphi_{\boldsymbol{y}}\rangle$ に漸化式 (10) を使おう．やはり $\boldsymbol{x}_1\neq\boldsymbol{y}$ なのでクロネッカーデルタの項は消えて，

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = \left(\frac{\zeta}{2d}\right)^2\sum_{\boldsymbol{x}_1\in\mathcal{N}(\boldsymbol{x})}\sum_{\boldsymbol{x}_2\in\mathcal{N}(\boldsymbol{x}_1)}\langle\varphi_{\boldsymbol{x}_2}\varphi_{\boldsymbol{y}}\rangle \tag{12}$$

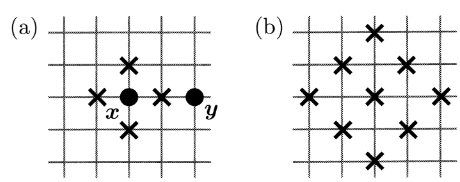


図 2　漸化式 (10) からランダムウォークが作られる様子．(a) 黒丸が格子点 $\boldsymbol{x}$ と $\boldsymbol{y}$ を表す．× 印は (11) で $\boldsymbol{x}_1$ が動く範囲（つまり $\boldsymbol{x}$ と隣り合う点）．(b) (12) で $\boldsymbol{x}_2$ が動く範囲．中央の × が $\boldsymbol{x}$ であり，右端の × が $\boldsymbol{y}$ である．

*4）　L^d 重積分の部分積分など知らないと思うかも知れないが，$\varphi_{\boldsymbol{x}}$ 以外の積分変数はすべて固定して，一変数の場合の部分積分の公式を使えばよい．

となる．$\boldsymbol{x}_1$ は $\boldsymbol{x}$ に隣り合う点を動き，$\boldsymbol{x}_2$ は各々の $\boldsymbol{x}_1$ に隣り合う点を動く．$d=2$ なら，$\boldsymbol{x}_2$ は図 2 (b) に × 印で示した 9 つの点の上を動く*5)．

当然，次は (12) 右辺の $\langle\varphi_{\boldsymbol{x}_2}\varphi_{\boldsymbol{y}}\rangle$ に漸化式 (10) を使うわけだが，ここで今までとは少し事情が変わる．(12) での和で，一通りだけ $\boldsymbol{x}_2$ が $\boldsymbol{y}$ に一致することがある．このときは $\delta_{\boldsymbol{x}_2,\boldsymbol{y}}$ が 1 になるので

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\left(\frac{\zeta}{2d}\right)^2 + \sum_{\boldsymbol{x}_1\in\mathcal{N}(\boldsymbol{x})}\sum_{\boldsymbol{x}_2\in\mathcal{N}(\boldsymbol{x}_1)}\sum_{\boldsymbol{x}_3\in\mathcal{N}(\boldsymbol{x}_2)}\left(\frac{\zeta}{2d}\right)^3\langle\varphi_{\boldsymbol{x}_3}\varphi_{\boldsymbol{y}}\rangle \tag{13}$$

という関係が得られる．

(12) に登場した $\omega:=(\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2)$ という格子点の並びに注目しよう．$\boldsymbol{x}_1$ が $\boldsymbol{x}$ の隣の点で，$\boldsymbol{x}_2$ が $\boldsymbol{x}_1$ の隣の点だったので，これを「はじめ $\boldsymbol{x}$ にあったものが，次に隣の格子点 $\boldsymbol{x}_1$ に移動し，それからそのまた隣の格子点 $\boldsymbol{x}_2$ に移動した」という風に見ることにしよう．このようにして，あるものが次々と隣の格子点に移動していくことで作られる格子点の並びを，**ランダムウォーク**と呼ぶ．移動していくものは「ランダムウォーカー」である．

各々の格子点に隣り合う格子点はちょうど $2d$ 個ずつあるから，$\boldsymbol{x}_1$ の選び方は $2d$ 通りあり，$\boldsymbol{x}_2$ の選び方は $2d$ 通りある．よって，2 ステップの（つまり，移動が 2 回の）ランダムウォーク $\omega:=(\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2)$ は全部で $(2d)^2$ 通りあることがわかる．(12) の右辺では，これら $(2d)^2$ 通りのランダムウォークがすべて足しあげられている．

これら $(2d)^2$ 個のランダムウォークのうち，(2 回とも $\boldsymbol{y}$ のいる方向に動いた) 一つだけが $\boldsymbol{x}_2=\boldsymbol{y}$ を満たす．これが，(13) の右辺の第一項に対応する．また，(13) の右辺第二項では，3 ステップのランダムウォーク $\omega:=(\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$ すべてが足しあげられている．

これから先の一般化を念頭において，(13) を

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\sum_{\substack{\omega:\boldsymbol{x}\to\boldsymbol{y}\\(|\omega|<3)}}\left(\frac{\zeta}{2d}\right)^{|\omega|} + \sum_{\substack{\omega:\boldsymbol{x}\to\cdot\\(|\omega|=3)}}\left(\frac{\zeta}{2d}\right)^3\langle\varphi_{\boldsymbol{x}_3}\varphi_{\boldsymbol{y}}\rangle \tag{14}$$

と書き直しておこう．もちろん説明が必要だ．

一般のランダムウォーク $\omega=(\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n)$ とは，$\{\boldsymbol{x}_{j-1},\boldsymbol{x}_j\}\in\mathcal{B}$ を満たす（つまり，番号が一つ違うところは隣り合っている）$n+1$ 個の格子点の組のことを言う．このランダムウォークは格子点 $\boldsymbol{x}_0$ から出発し $\boldsymbol{x}_n$ で終わる．ランダムウォークのステップ数を $|\omega|=n$ と書く．(14) 右辺の一つ目の和では，$\boldsymbol{x}$ から出発して $\boldsymbol{y}$ で終わるステップが 3 未満のランダムウォークすべてについて足しあげている（そのような ω は一通りで，$|\omega|=2$）．二つ目の和では，$\boldsymbol{x}$ から出発して勝手な点で終わるステップが 3 のランダムウォークすべてについて足しあげている．確かに (13) と同じである．

さて，(11), (12), (13) と進んできた路線で，基本の漸化式 (10) を n 回くり返し用いると，

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\sum_{\substack{\omega:\boldsymbol{x}\to\boldsymbol{y}\\(|\omega|<n)}}\left(\frac{\zeta}{2d}\right)^{|\omega|} + \sum_{\substack{\omega:\boldsymbol{x}\to\cdot\\(|\omega|=n)}}\left(\frac{\zeta}{2d}\right)^n\langle\varphi_{\boldsymbol{x}_n}\varphi_{\boldsymbol{y}}\rangle \tag{15}$$

という関係が得られる．ここで右辺の一つ目の和ではステップが n 未満のランダムウォークについて足しあげていることに注意．

ここで $n\uparrow\infty$ の極限を考える．ある定数 C があって $\langle\varphi_{\boldsymbol{x}_n}\varphi_{\boldsymbol{y}}\rangle\le C$ であること*6)，ステップ数が n のランダムウォークは $(2d)^n$ 通りあることを使うと，(15) 右辺の二つ目の和は，$(2d)^n\{\zeta/(2d)\}^nC=C\zeta^n$ 以下であることがわかる．よって，$n\uparrow\infty$ で，(15) 右辺の二つ目の和はゼロに収束し，(15) は

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\sum_{\omega:\boldsymbol{x}\to\boldsymbol{y}}\left(\frac{\zeta}{2d}\right)^{|\omega|} \tag{16}$$

*5) $\boldsymbol{x}$ に隣り合う点は 4 つあり，その各々に 4 つの隣り合う点があるから，和は 16 通りである．9 つの点のうちの多くは重複して和に登場することに注意．

*6) 任意の $\boldsymbol{x}$, $\boldsymbol{y}$ について $\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle$ が有限なので，その最大値を C と書いただけ．

という簡単な形になる．二点相関関数 $\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle$ はランダムウォークの和を使って正確に表現できるのだ．

6. 確率過程としてのランダムウォーク

ここで調和振動子の統計力学から離れて，確率過程としてのランダムウォークを定義しよう[2)]．以下のような確率的なルールで Λ 上の格子点の上を動き回るランダムウォーカーを考える．

(0) ランダムウォーカーは，はじめ格子点 $\boldsymbol{x}_0$ にいる．

(i) ランダムウォーカーは確率 $1-\zeta$ で消滅する（消滅したらそこで「試合終了」）．

(ii) 消滅しなかったとき，ランダムウォーカーは現在の格子点に隣り合う $2d$ 個の格子点のいずれかに等確率で移動する．(i) に戻る．

このルールに従ってランダムウォーカーを動かしていくと，格子上にウォーカーが動いた軌跡が描き出される．この軌跡を**ランダムウォーク**と呼ぶ．ランダムウォークは，格子上の隣り合う点を次々と線で結んだようなものとイメージしよう．ただし，ウォーカーは同じ格子点を何度も通ってよいから，一般にランダムウォークは自分自身と何度も交わる複雑な図形になる（図 3）．

このランダムウォークは前節に登場したものと同じである．実際，上のルールを 1 回使うと，ランダムウォーカーは確率 ζ で生き残り，さらに，確率 $1/(2d)$ で隣の格子点のいずれかに移動する．よって，1 ステップの後には，確率 $\zeta/(2d)$ で $\boldsymbol{x}_0$ の隣の格子点のいずれかにいることになる．同様に考えれば，n ステップ後には確率 $\{\zeta/(2d)\}^n$ で，$\boldsymbol{x}_0$ から出発した $(2d)^n$ 通りのランダムウォークの内のいずれかを辿っていることになる．これは，(16) の和に現れた重みそのものである．

これによって，二点相関関数の表現 (16) をより「確率論らしい」形に書き直すことができる．(16) では $\boldsymbol{x}$ から $\boldsymbol{y}$ に向かう様々なステップ数のランダムウォークを足しあげているわけだが，これは，

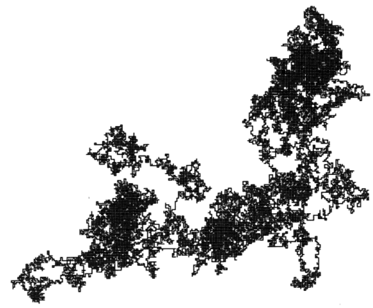


図 3　50000 ステップのランダムウォークの例．これは次元 2 のフラクタル図形であることが知られている．

$\boldsymbol{x}$ から出発した（終点を制限しない）一つのランダムウォークを考えて，それが $\boldsymbol{y}$ を通過するたびにカウントするのと同じことだ．そこで，$\boldsymbol{x}$ から出発したランダムウォークが $\boldsymbol{y}$ を通る回数を $N_{\boldsymbol{x}\to\boldsymbol{y}}$ と書けば，統計力学の二点相関関数を

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle \tag{17}$$

と表現できることになる．

7. ランダムウォーク表現の応用

(17) を使えば，二点相関関数のいくつかの性質を（積分を実行することなく）知ることができる．先に進む前に，ランダムウォーク ω のステップ数 $|\omega|$ がちょうど n に等しい確率は $p_n = \zeta^n(1-\zeta)$ であることに注意しておく（$n=0,1,2,\ldots$）．

$\boldsymbol{x}$ から $\boldsymbol{y}$ に向かうランダムウォークの最小のステップ数を $D(\boldsymbol{x},\boldsymbol{y})$ と書こう．明らかに，$\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle$ はステップ数が $D(\boldsymbol{x},\boldsymbol{y})$ 以上のランダムウォークが訪れる点の総数よりも小さいから，

$$\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle \le g_0 \sum_{n=D(\boldsymbol{x},\boldsymbol{y})} n\,p_n$$
$$= g_0\left(D(\boldsymbol{x},\boldsymbol{y})+\frac{\zeta}{1-\zeta}\right)\zeta^{D(\boldsymbol{x},\boldsymbol{y})} \tag{18}$$

という上界が得られる．$0\le\zeta<1$ である限り，二点相関関数 $\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle$ は（グラフ理論的）距離 $D(\boldsymbol{x},\boldsymbol{y})$ について指数的に減衰することがわかる．

また，$\sum_{\boldsymbol{y}\in\Lambda} N_{\boldsymbol{x}\to\boldsymbol{y}}$ はランダムウォークの総ステップ数そのものであることに注意すると，

$$\sum_{\boldsymbol{y}\in\Lambda}\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle = g_0\sum_{n=0}^{\infty} n\,p_n = g_0\frac{\zeta}{1-\zeta} \qquad (19)$$

のように，二点相関関数の総和を簡単に求めることができる．

次の応用として，$\zeta=1$（つまり，$\mu:=\kappa/(k_{\mathrm{B}}T)=0$）とした系が定義できるかという問題を検討しよう．

$\zeta=1$ ならランダムウォーカーは決して消滅しないので，格子の中を永遠にランダムに動き回る．明らかに，どのような $\boldsymbol{y}$ についても $\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle=\infty$ であり，(17) より，二点相関関数も無限大になる．

ところが，格子が無限に多いと，話はそう単純ではない．まず $\zeta<1$ として問題をきちんと定義し，それから，$L\uparrow\infty$ の極限をとった後で，$\zeta\uparrow 1$ とすることを考えてみよう．

数学的にはやや厳密さを欠くが，この場合には，無限の格子の中を永遠に動き回るランダムウォーカーを考えればよいことになる．格子点が無限にあるから，$\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle$ が無限になるとは限らない．

実際，これはランダムウォークの再帰性という重要な性質に直結している．格子の次元が $d=1,2$ のときには，ランダムウォークは再帰的であり，（格子が無限に大きくても）何度でも出発点に戻ってくることが知られている．よって，$\lim_{\zeta\uparrow 1}\lim_{L\uparrow\infty}\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle=\infty$ であり，$\zeta\uparrow 1$ の系に意味を付けることはできない．一方，次元が 3 以上ではランダムウォークは非再帰的であり，いずれは格子のはるか遠い領域へと逃げていってしまうことが知られている．そのため，$\lim_{\zeta\uparrow 1}\lim_{L\uparrow\infty}\langle N_{\boldsymbol{x}\to\boldsymbol{y}}\rangle<\infty$ となる．つまり，$\zeta\uparrow 1$ の系をしっかりと定義することができる！このように格子の次元によって，系の性質が本質的に異なっているのは面白いところだ[*7]．また，この極限での二点相関関数は，$|\boldsymbol{x}-\boldsymbol{y}|$ が大きいところで

$$\lim_{\zeta\uparrow 1}\lim_{L\uparrow\infty}\langle\varphi_{\boldsymbol{x}}\varphi_{\boldsymbol{y}}\rangle\sim\frac{\text{定数}}{|\boldsymbol{x}-\boldsymbol{y}|^{d-2}} \qquad (20)$$

のように振る舞うことが知られている．これはランダムウォークの軌跡が次元 2 のフラクタル図形であることの反映である[*8]．

8. そして，その先の話

バネが絡み合った系の統計力学と，格子上をデタラメに動き回るランダムウォークという，一見すると何の関連もなさそうな問題が密接に関わっていることを見てきた．言うまでもなく，これは統計力学としても，確率論としても，多くのことが知られている初等的な例題である．しかし，より非自明な問題においても，統計力学モデルと（ランダムウォークやパーコレーションに代表される）確率幾何的な問題が結びつくことは珍しくない[*9]．そのような対応関係をうまく活かせば，統計力学（や場の理論）と確率論の双方の手法を利用できるようになるだけでなく，困難な問題を異なった角度から見ることで，より豊かな物理的・数理的直観を養うことができる．現代の統計力学は物理と数学が出会うことで新たな手法・概念・視点が生み出される重要な発展の場なのである[3]．

参考文献

1) 田崎晴明『統計力学』（培風館）．
2) 1 次元でのランダムウォークについては，池田信行他『確率論入門 I』（培風館）が詳しい．
3) 関連する問題の発展については，近刊の田崎晴明，原隆『相転移と臨界現象の数理』（共立出版）を見よ．

（たざき・はるあき，学習院大学理学部）

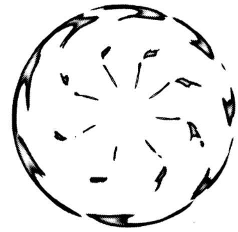
---

*7）この性質は，統計力学では「1，2 次元では結晶は安定に存在しない」という事実に対応し，場の量子論では「時空間 1，2 次元の質量 0 の自由スカラー場は存在しない」という事実に対応する．

*8）$R=|\boldsymbol{x}-\boldsymbol{y}|$ とする．$\boldsymbol{x}$ を中心にした半径 R の球面上の格子点の数は R^{d-1} 程度．$\boldsymbol{x}$ を含む次元 ν の図形が球面と交わる点の数は $R^{\nu-1}$ 程度．よって，この図形が $\boldsymbol{y}$ を訪れる回数の期待値は $R^{\nu-1}/R^{d-1}=R^{-(d-\nu)}$ 程度．

*9）例えば，ポテンシャル (3) に新たに $\lambda\sum_{x\in\Lambda}(\varphi_{\boldsymbol{x}})^4$ という非線形の項を付け加えた系は φ^4 モデルと呼ばれており，スピン系の相転移や相互作用する場の量子論の基本的なモデルである．ここで扱ったモデルと似ているが，本質的に難しく，はるかに豊かな性質を持っていると期待されている．このモデルを扱う際にも，やはりランダムウォーク（ただし，自分自身や他のランダムウォークと相互作用するようなランダムウォーク）との対応は重要な役割を果たす．

ランダム行列とその周辺



香取　眞理

1. はじめに

コインを何度も投げて表の出る回数の分布を調べることにする．表の出る確率が p，裏の出る確率が $q=1-p\ (0<p<1)$ と毎回決まっているとしたときには，N 回のコイン投げのうち表の出る回数 X がちょうど n 回である確率は

$$P_N(X=n)={}_NC_n p^n q^{N-n} \tag{1}$$

で与えられる．ここで ${}_NC_n$ は二項係数 $N!/\{n!(N-n)!\}$ であり，この分布を**二項分布**と言う．表の出る回数 X は**確率変数**，あるいは値がランダムであるということから**乱数**と呼ばれる．

さて，この二項分布の式からどのような話が展開できるであろうか．ちょっと想像してもらいたい．もちろん (1) 式から平均値 $\mu=\langle X\rangle=Np$ や分散 $\sigma^2=\langle(X-\mu)^2\rangle=Npq$ を計算しながら確率論入門の講義を進めるのは常套手段であろう．適当な条件の下で試行回数 $N\to\infty$ の極限をとると，二項分布の極限としてポアソン分布が得られたり（小数の法則），ガウス分布（正規分布）が得られたり（大数の法則）する話は**極限定理**の説明に最適である．

ところで，コインは金属でできた剛体であるから，空中に投げられたコインが手のひらや机の上に表と裏のどちらの面を上にして着地するかは**力学の問題**であるはずである．本来なら，コインを投げ上げたときの位置と初速度を初期条件として与えてニュートンの運動方程式を解くべきと思われる問題を，我々はどうして**確率の問題**に置き換えて考えることができるのであろうか．もしかすると，コイン投げの話から**カオス現象**を議論することもできるのかもしれない．

あるいは，コインという物体のことは忘れて，表と裏という 2 つの状態のいずれかをとるという点に注目すると，全く別の話の展開も考えられる．原子や分子といったミクロな実体は**量子力学**で記述される．量子力学の世界では，エネルギーの値がとびとびになるだけでなく，回転の向きと強さを表す角運動量ベクトルも離散化される．電子は原子の中で原子核の周りを「公転」しているだけでなく「自転」もしている．一般に素粒子の自転を**スピン**と呼ぶが，電子の持つスピンは最小のものであり，回転軸を指定するとその周りに大きさ $\hbar/2$ で右回りに回転するか，同じ大きさ $\hbar/2$ で左回りに回転するかしかあり得ない．ここで $\hbar$ は $h=6.62\times10^{-34}$ [Js] という微小ながらも零ではない定数（プランク定数）を 2π で割った値である．電子のスピンが右向き回転している状態をコインの表に，左向き回転している状態をコインの裏にそれぞれ対応させることにしよう．すると，コイン投げの問題は電子のスピン状態を何度も観測する問題に発展する．ファインマンはこのスピンの話題から有名な量子力学の講義を始めている[1)]．

二項分布と熱平衡状態にある二準位系のカノニカル分布（正準分布）とを比較して，**統計力学**を

解説することもできるだろう[2]．コインの表裏の出方を時系列として見れば**ランダム・ウォーク**という**確率過程**を導入することができるし，その連続極限として**ブラウン運動**を定義することも可能である．

さて，本稿の主題は**ランダム行列**である．ランダム行列とは「成分が乱数である行列」のことである．たった一つの乱数 X を考えても，上述のように話はどんどんと広がっていく．ランダム行列の話題の広がり方は想像を絶するものがある．そしてその最大の魅力は，まさに現在，多くの研究者によってその拡張が繰り広げられているということである[3,4]．

2. 何を表す行列がランダムなのか

各成分がランダムであるベクトルや乱数の列ではなく，**ランダムな行列**を考えるところが重要である．行列は，数値を2次元的に配置した数表であり，ベクトルの変換を表し，線形連立方程式の係数のセットを与え，あるいはグラフの頂点（集合の要素）の間の連結の仕方（相互作用）を表すことができる．ハイゼンベルクの不確定性原理とは量子力学的な実体の位置と運動量の値を同時に決定することはできないというものであるが，これは位置を測定してから運動量を測定した場合と，その逆にまず運動量を測定してから次に位置を測定した場合とでは結果が違うということである．行列の積は一般に非可換であり，量子力学を表現するのに行列は不可欠と言える．また，行列を使って系の持つ対称性や群の構造を表現することができる．**行列に値を持つ乱数**は，こういった諸々のもの（変換，相互作用，ハミルトニアン，遷移則，対称性）の乱れや揺らぎを表すことになる．ランダム行列理論がどのような行列に応用されているか，具体例で見てみることしよう．

2.1 複数の時系列の数表としてのランダム行列

ランダム行列理論は**経済物理学**[5~7]に応用されている．いま，N 種の株式 $i=1,2,\ldots,N$ に対して株価の変動を記録した各々長さ M の時系列 $(\delta x_i(t_1),\delta x_i(t_2),\ldots,\delta x_i(t_M))$ が与えられているとする．このデータは N 行 M 列の数表，すなわち $N\times M$ の実長方行列

$$W=\begin{pmatrix}\delta x_1(t_1) & \ldots & \delta x_1(t_M)\\ \delta x_2(t_1) & \ldots & \delta x_2(t_M)\\ \vdots & \ddots & \vdots\\ \delta x_N(t_1) & \ldots & \delta x_N(t_M)\end{pmatrix} \quad (2)$$

で表される．この行列 W の転置である $M\times N$ 実行列を ${}^tW$ とおいて

$$C=\frac{1}{M}W\,{}^tW \quad (3)$$

という $N\times N$ の正方行列を定義する．いま仮に，株価の変動 $\delta x_i(t_j)$ が株の銘柄 i にも取引時刻 t_j にも依存せず，ある分布関数に従う独立な乱数であるとしよう．本稿の冒頭で述べたコイン投げの話で，表と裏の出る確率が毎回 p と $q=1-p$ と定まっていたようにである．ランダム行列理論はこのような場合に，正方行列 C の固有値と固有ベクトルの分布に対して正確な答えを与えることができる．それでは，実データ (2) から直接得られる行列 C の固有値と固有ベクトルを数値計算してみたら，その分布はどうなっているのであろうか．結果は一般には，ランダム行列理論で計算したものからずれる．行列 C は株価変動における相関を定量的に表現する行列であり**相関行列**と呼ばれている．その実データから，完全にランダムであるとしたときの分をうまく差し引くことができれば，市場の動向を捉えることができよう．相関行列の解析データは，リスクの低いポートフォリオを構築するのに役立つはずである[7]．

2.2 変換行列としてのランダム行列

次に**無線通信理論**におけるランダム行列の応用例を見ることにしよう[8,9]．いま，M 人の携帯電話ユーザーが別々の情報を同一の基地局に向けて送信しようとしている．受け手の基地局で各ユーザーからの情報が分離できるようにするには，どのようにして送信すればよいだろうか．各ユーザーが送信するデータを $x_j, 1\leq j\leq M$ とする．分離可能とするために，各々を別々の**変調信号**に乗せて送信

することが行われている．各変調信号が N 次元列ベクトル $\boldsymbol{a}_j = {}^t(a_{1j}, a_{2j}, \ldots, a_{Nj}), 1 \le j \le M$ で表されるものとすると，j 番目のユーザーは $\boldsymbol{a}_j x_j$ という長さ N の変調データを送信することになる．基地局では，全 M ユーザーからの変調信号の重ね合わせに，さらに通信路ノイズ（これは $\boldsymbol{n} = {}^t(n_1, n_2, \ldots, n_N)$ で表すことにする）が加えられた信号が受信される．M ユーザーからのデータをまとめて M 次元の入力ベクトル $\boldsymbol{x} = {}^t(x_1, \ldots, x_M)$ で表し，基地局での受信信号を N 次元の出力ベクトル $\boldsymbol{y} = {}^t(y_1, \ldots, y_N)$ で表すと

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{n} \tag{4}$$

という関係が成り立つことになる．ただし，A は各 j 列が列ベクトル $\boldsymbol{a}_j$ で与えられる $N \times M$ 実行列である．

この**線形ベクトル通信モデル** (4) は，実際に携帯電話などで使われている通信路（CDMA 通信路）のモデルと見なすことができる．変調信号 $\boldsymbol{a}_j$ は拡散符号系列と呼ばれており，実際には擬似乱数系列が使われているそうである[8]．(4) 式の A は M 次元ベクトル $\boldsymbol{x}$ から N 次元ベクトル $\boldsymbol{y}$ への変換を表す行列であるが，それをランダム行列で近似することができるのである．

任意の $N \times M$ 実行列 A は，それぞれ $N \times N$ と $M \times M$ である 2 つのサイズの異なるユニタリ行列 U と V，および非対角成分はすべて零であり，対角成分は非負実数であるような $N \times M$ 行列 D を用いて

$$A = UDV^\dagger \tag{5}$$

と分解できる．ただし，$V^\dagger$ は V の共役転置（エルミート共役）行列である．$L = \min\{N, M\}$ とすると，D の対角成分は L 個あるが，これらを $d_i, 1 \le i \le L$ と書くことにする．すると

$$\begin{aligned} AA^\dagger &= UD\,{}^tDU^\dagger, \\ D\,{}^tD &= \mathrm{diag}(d_1^2, d_2^2, \ldots, d_L^2) \end{aligned} \tag{6}$$

であるから，行列 $AA^\dagger$ の固有値は $\lambda_i = d_i^2, 1 \le i \le L$ で与えられることになる．

通信路の情報伝達能力を表す量に相互情報量という指標がある．これは上の固有値 $\lambda_i, 1 \le i \le L$ で表すことができる．A の各成分が独立同分布であると仮定すると，上述のようにランダム行列理論は $AA^\dagger$ の固有値分布を正確に与えることができる．これにより，相互情報量の理想値（通信路容量）が計算できるのである[8,9]．

2.3 連立方程式の係数行列としてのランダム行列

生態系では様々な種類の生物が種間相互作用の結果，それぞれの種の個体数を変化させている．いま，N 種から成る生態群集を考え，そのうちの i 番目の種の時刻 t での個体数を $x_i(t)$ とする．この値の時間変化が

$$\frac{dx_i(t)}{dt} = f_i(x_1, \ldots, x_N), \quad 1 \le i \le N \tag{7}$$

という連立方程式に従っているというモデルを考える．右辺の関数は，種間の複雑な相互作用を表すものであるから，一般に (7) は非線形方程式系となる．その**平衡解**，すなわち $dx_i^0/dt = f_i(x_1^0, \ldots, x_N^0) = 0, 1 \le i \le N$ となる解 $\boldsymbol{x}^0 = {}^t(x_1^0, \ldots, x_N^0)$ も多数あり得る．このような場合，各々の平衡解の**安定性**を調べて比較することが重要である．長時間後には，系は最も安定な解で表される状態に落ち着くことが期待されるからである．$\boldsymbol{x} = {}^t(x_1, \ldots, x_N)$ として，一つの平衡解 $\boldsymbol{x}^0$ からのずれを $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{x}^0$ と置くことにする．$\boldsymbol{y}$ について (7) を線形化すると，$M = (m_{ij}) = (\partial f_i / \partial x_j |_{\boldsymbol{x}=\boldsymbol{x}^0})$ を $N \times N$ のヤコビ行列として

$$\frac{d\boldsymbol{y}(t)}{dt} = M\boldsymbol{y}(t) \tag{8}$$

という N 連立線形微分方程式が得られる．

係数行列 M の対角成分 m_{ii} は負であろう．種 i の個体数が平衡値を超えれば（$y_i > 0$）食料が不足するので，その種の個体数は減る（$dy_i/dt < 0$）．逆に $y_i < 0$ なら食料に余裕があり，その種は増えることができる（$dy_i/dt > 0$）．したがって，$dy_i/dt = m_{ii} y_i + \cdots$ としたとき $m_{ii} < 0$ である．それでは M の非対角成分 $m_{ij}, i \ne j$ はどうであろうか．種 j が種 i の餌である場合，種 j の

個体数が普段の値（平衡値）より増加すると，それを捕食している種 i の個体数も増えるだろうから $m_{ij}>0$ である．反対に，種 j が種 i の天敵である場合は，種 j が普段より増えると種 i は多く食われてしまうので減少するから $m_{ij}<0$ である．このように，種 i と種 j とが**捕食・被食関係**にある場合は，m_{ij} と m_{ji} とは負の相関（$m_{ij}m_{ji}<0$）を持つことになる．これに対して，種 i と種 j が**共生関係**にあるときには，正の相関（$m_{ij}m_{ji}>0$）を持つことになる．

一般に，どのような種間相互作用が支配的なときに系は安定になるのであろうか．**数理生態学**[10]という研究分野では，特定の生態系ではなく，大規模生態系一般の特性を抽出したいと考える．このような目的で，係数行列 M をランダム行列として扱い，系の振舞いを統計的に記述しようとする試みが盛んである[11,12]．この場合には，行列 M の成分を全く無相関にランダムにすることはしない．対角成分は負の値に定める．また，非対角成分の対 m_{ij} と m_{ji} $(i\neq j)$ との間には相関を持たせる．例えば $\langle m_{ij}m_{ji}\rangle=\gamma$ としたければ，$\gamma<0$ に対しては行列 M を**実反対称行列**，$\gamma>0$ に対しては**実対称行列**として，上半非対角成分 $m_{ij}, i<j$ を平均零，分散 γ で分布させればよいだろう．前者は捕食・被食関係が支配的な生態系を，後者は共生関係が支配的な生態系をそれぞれモデル化したことになる．行列 M の最大固有値が正なら，平衡値からのずれは時間とともに増大するので，系は不安定，負なら系は安定（零なら局所安定）である．ランダム行列理論を用いて固有値分布を計算することにより，大規模生態系の時間発展を予測することができる．

グラフ理論では**隣接行列** $G=(g_{ij})$ が系の代数的表現に便利である．これは，グラフ上の2つの頂点 i と j が隣接していて辺で結ばれているときだけ $g_{ij}=1$ であり，それ以外では $g_{ij}=0$ としたものである．2つの頂点 i と j を結ぶ辺を，i から j に至る辺とその逆向きの j から i に至る辺とに分けて考え，i から j への辺は g_{ij} 本あるのに対して，逆向きの $j\to i$ なる辺は g_{ji} 本であり，$g_{ij}\neq g_{ji}$ であるようなグラフを考えることもできる（**有向グラフ**）．さらには，そもそもは隣接していなかった頂点間 (i,k) でも $g_{ik}\neq 0$ とすれば，遠く離れた頂点どうしも絡み合う**複雑ネットワーク**のモデルとなる[13~17]．連立方程式の係数行列はネットワークにおける頂点間の相互作用を表すものと見なせる．複雑ネットワークの研究にランダム行列理論は有効なのであろうか．これまでのランダム行列理論では，$\langle g_{ij}g_{ji}\rangle$ といった非対角成分の対の間の相関は考慮するが，3つ以上の成分の間の相関はあまり考えられてない．しかし，上述の生態系では**食物連鎖**は重要であり，これは $\langle g_{ij}g_{jk}g_{ki}\rangle\neq 0$ といった相互作用の高次相関で表されるものである[11]．グラフ理論で言えばこれは**ループ構造**に他ならない．話題はどんどんと広がっていく．

3. 量子物理学とランダム行列

「系が今ある状態 i にあるものとしよう.」このように設定して話を始めるのは構わないが，系が量子力学に従う場合は気をつけてフォローしなければならない．ハイゼンベルクの不確定性原理があるので，一般に系は一つの状態に留まってはくれない．別の状態に遷移していくのである．一つの状態 i から別の状態 j へ遷移する度合いは複素数 H_{ij} で与えられる．これにその複素共役 $\overline{H}_{ij}$ をかけると非負の実数 $|H_{ij}|^2\geq 0$ が得られるが，これが状態 i から状態 j に遷移する確率を与える．H_{ij} は**遷移確率振幅**と呼ばれる．

このように量子力学的実体はふらふらとその状態を変化させるものであるから，そもそも「今ある状態 i にある」としたのには無理があった．時刻 t に状態 i にある確率振幅 $\psi_i(t)$ を指定して，その状態にある確率 $|\psi_i(t)|^2$ を与えることしか我々にはできない．系のとり得る状態の数が N 個とすると，各時刻 t での量子状態は N 成分複素ベクトル $\boldsymbol{\psi}(t)={}^t(\psi_1(t),\ldots,\psi_N(t))$ で指定され，その時間発展は

$$\sqrt{-1}\hbar\frac{d\boldsymbol{\psi}(t)}{dt} = H\boldsymbol{\psi}(t) \tag{9}$$

で与えられる．ここで，$\hbar = h/(2\pi)$，また $H = (H_{ij})_{1\le i,j\le N}$ は $N\times N$ の複素正方行列であり，**ハミルトニアン**（**ハミルトン行列**）と呼ばれる．これが**シュレーディンガー方程式**である．

この式を (8) 式と見比べると，ハミルトニアンは連立方程式の係数行列の一種であることが分かる．これは，状態 i と状態 j との相関を表す量である．ただし量子力学ではこの相関は正負や大きさを表す実数ではなく，大きさと位相を表す複素数 H_{ij} で表されるのである．そして重要なことは，シュレーディンガー方程式によって，確率の和は 1 に保存されなければならないということである．つまり，任意の Δt に対して $\sum_{i=1}^{N}|\psi_i(t+\Delta t)|^2 = \sum_{i=1}^{N}|\psi_i(t)|^2 = 1$ である．この要請から，H は**エルミート行列**でなければならないことになる：$H^\dagger = H$, すなわち，

$$\overline{H}_{ji} = H_{ij}, \quad 1\le i,j\le N. \tag{10}$$

エルミート行列はユニタリ行列 U で対角化される：$UHU^\dagger = \mathrm{diag}(E_1, E_2, \ldots, E_N)$．固有値 $E_i, 1\le i\le N$ はすべて実数であり，この系のとり得るエネルギーの値を与える．これらを**エネルギー準位**と言う．(9) 式の両辺に左から U を掛けると，U のユニタリ性 $U^\dagger U = I$（単位行列）から $\sqrt{-1}\hbar d(U\boldsymbol{\psi}(t))/dt = (UHU^\dagger)(U\boldsymbol{\psi}(t))$ となる（H と U は時間に依存しないものとした）．$\boldsymbol{\varphi}(t) \equiv U\boldsymbol{\psi}(t) = {}^t(\varphi_1(t), \ldots, \varphi_N(t))$ と置くと，(9) 式は独立な N 個の式

$$\sqrt{-1}\hbar\frac{d\varphi_i(t)}{dt} = E_i\varphi_i(t), \quad 1\le i\le N \tag{11}$$

に分離される．$\varphi_i(t)$ はエネルギー準位 E_i の固有状態であり，これは $\psi_j(t)$ の重ね合わせ $\varphi_i(t) = \sum_{j=1}^{N} U_{ij}\psi_j(t)$ で与えられる．

量子力学の重要な視点は，(11) よりも (9) を本質と見ることである．エネルギー固有状態 φ_i に対してのみ成り立つ式 (11) ではなく，それをユニタリ変換したすべての状態 $\boldsymbol{\psi}(t) = U^\dagger\boldsymbol{\varphi}(t)$ に対する式 (9) が大切なのである．そもそも量子力学とは不

確定性原理を表現する力学だからである．H と ψ は連動して変換される（$H \to UHU^\dagger, \psi \to U\psi$）．しかしこの変換は表示を変えるだけであり，記述される対象であるところの量子系自体は不変なのである．

ランダム行列理論は，多数の核子（陽子と中性子）からなる原子核の複雑な量子状態を，統計的に記述する方法として物理学に導入された[18]．そこでは，ハミルトニアン H がランダム行列となる．よって，エルミート性 (10) は大切に保持されなければならない．ユニタリ変換で量子系は不変であるので，H の出現確率 $P(H)dH$ も**ユニタリ不変**でなければならない．

一般に，行列 H が正則な行列 V で相似変換 $H \to VHV^{-1}$ されるとき，この変換で不変な量はすべて H のべき乗のトレース $\mathrm{tr}H^p$ で与えられる．ユニタリ変換は相似変換の一例である．さて，$N \times N$ の複素行列 H にはエルミート条件 (10) が課されているので，自由度は N^2 だけである．実際，N 個の実対角成分 $H_{11}^{\mathrm{R}}, \ldots, H_{NN}^{\mathrm{R}}$ と $N(N-1)/2$ 個の上半非対角成分 $H_{ij}, 1 \leq i < j \leq N$ の実部 $H_{ij}^{\mathrm{R}} = \mathrm{Re}(H_{ij})$ と虚部 $H_{ij}^{\mathrm{I}} = \mathrm{Im}(H_{ij})$ の合計 $N + N(N-1)/2 \times 2 = N^2$ 個の実数を与えると，$N \times N$ のエルミート行列 H が定まる．ランダム行列の統計集団として，これら N^2 個の成分が独立に分布する場合を考えることにしよう．冒頭で述べたコイン投げの話で，毎回の試行結果を独立に分布するとしたのと同じようにである．エルミート行列全体を表す空間の体積要素を dH と書き，ランダムなエルミート行列の確率密度を $P(H)$ と記すと，上述の独立性は，ある関数のセット $\{f_{ij}^{\#} : \# \in \{\mathrm{R}, \mathrm{I}\}, i \leq j\}$ があって，$P(H)dH$ は積

$$\prod_{1 \leq i \leq j \leq N} f_{ij}^{\mathrm{R}}(H_{ij}^{\mathrm{R}}) dH_{ij}^{\mathrm{R}} \prod_{1 \leq k < l \leq N} f_{kl}^{\mathrm{I}}(H_{kl}^{\mathrm{I}}) dH_{kl}^{\mathrm{I}}$$

と書けることを意味する．こうしておいて，確率 $P(H)dH$ が H に対する任意のユニタリ変換で不変であるという条件を課すと，関数 $\log f_{ij}^{\#}(x)$ は x の 2 次式 $ax^2 + bx + c$ でなければならないことが導かれる．そもそもユニタリ変換不変な量は $\mathrm{tr}H^p$ の関数で表されるという上述の一般論の結果と合わせると，行列 H の出現確率は

$$P(H)dH = \frac{1}{Z} e^{-a\,\mathrm{tr}\,H^2 + b\,\mathrm{tr}\,H} dH \qquad (12)$$

という形に定まってしまうのである[18]．ただし，$a > 0$ で b は実数，Z は規格化定数（分配関数）である．ランダム行列に量子力学からの要請としてエルミート条件を課す．また，確率分布として扱いやすくするために，自由度ごとの**独立性**を要請する．すると，独立な成分がいずれもガウス分布するランダム行列の統計集団が得られるのである．これを**ガウス型ユニタリ集団**（**Gaussian unitary ensemble, GUE**）と言う[3,18]．

4. おわりに

この世界は 1 g あたり 10^{23} 個という膨大な数（**アボガドロ数** $N_{\mathrm{A}} = 6.02 \times 10^{23}\,[\mathrm{mol}^{-1}]$）の原子や分子が集まってできている．それらの運動は大変複雑でありカオスである．統計力学では系のミクロな状態の時間変化を追うことはせず，エネルギーが等しい状態はどれも等しい確率で出現する（**等エネルギー面上の一様分布**）として物理量を計算する（**カノニカル分布**）．$\mathrm{tr}\,H^p = \sum_{j=1}^{N} E_j^p$ であるから，GUE の分布 (12) にある重み $e^{-a\,\mathrm{tr}\,H^2 + b\,\mathrm{tr}\,H}$ は統計力学での**ボルツマン因子** $e^{-\beta E}$ と同様に（エネルギー）固有値のみで値が決まる．エルミート行列 $H = U^\dagger \mathrm{diag}(E_1, \ldots, E_N) U$ のユニタリ成分 U については**一様分布**としているのである．これがユニタリ不変性の意味である．

「カオス的運動の時間平均を統計集団における平均で置き換えて計算する」という統計力学は，熱平衡系に対して成功をおさめた．多くの研究者はいま，**非平衡系のダイナミクス**に興味の中心をシフトしつつある．二項分布からランダム・ウォークやブラウン運動に至る道と同様の道筋が，ランダム行列とその周辺から四方に伸びている[19,20]．これらの道筋を逆に辿るとどこに行くのか．ランダム行列でよく記述できる現象はあまねく**カオス**

力学系を意味するのであろうか[21)].

参考文献

1) R.P. ファインマン，砂川重信訳：『ファインマン物理学 V 量子力学』，岩波書店 (1986).
2) 香取眞理，中野徹：連載「問題例で深める物理 第 3 回」数理科学 **550**, 78 (2009 年 4 月号)，サイエンス社.
3) 永尾太郎：『ランダム行列の基礎』，東京大学出版会 (2005).
4) 永尾太郎他：数理科学 **524**, 特集「ランダム行列の広がり」(2007 年 2 月号)，サイエンス社.
5) 高安秀樹他：数理科学，特集「エコノフィジックス最前線」(2002 年 10 月号)，サイエンス社.
6) J.-P. ブショー，M. ポッター，森平爽一郎監修，森谷博之，熊谷善彰訳：『金融リスクの理論—経済物理からのアプローチ—』，朝倉書店 (2003).
7) 相馬亘，藤原義久，尹熙元：経済物理とランダム行列，数理科学 **524**, 44 (2007 年 2 月号)，サイエンス社.
8) 田中利幸：無線通信におけるランダム行列，数理科学 **524**, 50 (2007 年 2 月号)，サイエンス社.
9) A.M. Tulino and S. Verdú, "Random Matrix Theory and Wireless Communications", Now Publishers (2004).
10) 巌佐庸：『生命の数理』，共立出版 (2008).
11) 時田恵一郎：生態学，数理科学 **524**, 56 (2007 年 2 月号)，サイエンス社.
12) A. Crisanti, G. Paladin, and A. Vulpiani: "Products of Random Matrices in Statistical Physics", Springer (1993).
13) 増田直紀，今野紀雄：『複雑ネットワークの科学』，産業図書 (2005).
14) 増田直紀，今野紀雄：『「複雑ネットワーク」とは何か』，講談社ブルーバックス (2006).
15) 今野紀雄，井手勇介：『複雑ネットワーク入門』，講談社サイエンティフィック (2008)
16) 今野紀雄，町田拓也：『図解入門よくわかる 複雑ネットワーク』，秀和システム (2008).
17) 青木秀明，相馬亘，藤原義久共編著，SGC ライブラリ『ネットワーク科学への招待』，サイエンス社 (2008).
18) M.L. Mehta, "Random Matrices", 3rd ed., Elsevier (2004).
19) 香取眞理他：数理科学 **546**, 特集「確率論的物理観」(2008 年 12 月号)，サイエンス社.
20) 香取眞理，種村秀紀：非衝突過程・行列値過程・行列式過程，数学 **61**, No.3 (2007) 225–247.
21) マーカス・デュ・ソートイ，冨永星訳：『素数の音楽』，新潮社 (2005).

(かとり・まこと，中央大学理工学部)

化学で重要な確率の概念

中田　宗隆

1. はじめに

数学と異なり，化学は身の回りに実際に存在する物質の構造や性質を調べる学問である．物質は有機化合物と無機化合物に分けることができる．有機化合物は，一部の化合物（一酸化炭素，二酸化炭素，炭素塩，シアン化物）を除く1個以上の炭素原子を含む化合物のことである．そして，無機化合物は有機化合物以外の化合物である．それぞれの化合物を対象として，**有機化学**と**無機化学**という学問分野がある（図1）．

化合物の構造や性質を研究する方法は二つある．一つは実験であり，いろいろな分析法が開発されている．このような学問分野を**分析化学**と呼ぶこともある．もう一つの方法は，どうしてそのような構造になるか，どうしてそのような性質になるかを理論的に解釈する学問分野である．以前は理論化学と呼ばれたこともあったが，現在では，**物理化学**と呼ばれることが多い．物理の理論に基づいて研究するからである．

それぞれの分野は独立しているわけではない．特に，分析化学と物理化学は密接に結びついている．分析によって，新しい事実が分かり，新しい事実から新しい理論が組み立てられ，新しい理論から新しい分析技術が生まれている．

物理化学をさらに大きく分けると，巨視的に研究する**熱力学**と，微視的に研究する**量子化学**がある．前者は古典力学（ニュートン力学），統計力学に基づいた学問体系であり，19世紀までにほぼ完成されたと言われている．一方，後者は量子力学に基づいた学問体系であり，20世紀になってから誕生し，現在も発展を続けている．そして，数学で重要な**確率**の概念は，熱力学でも量子化学でも重要な役割を果たしている（表1）．ここでは，熱

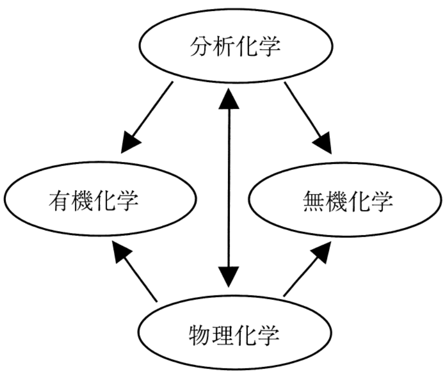


図1　化学のおもな四つの学問分野．

表1　物理化学の二つの大きな分野．

	熱力学	量子化学
歴史	19世紀までにほぼ完成	20世紀から発展
理論の基礎	古典力学	量子力学
ものの見方	巨視的な見方（Macroscopic）	微視的な見方（Microscopic）
扱う対象	気体，液体，固体，結晶など	核子，電子，分子，原子，電磁波など
確率の概念が重要な物理量	エントロピー変化など	波動関数，電子分布など

力学における**エントロピーの概念**と，量子化学における**波動関数の概念**を中心に紹介をする．

2. 二つの気体は自然に混ざる

化学は「化ける学問」とも読めるように，もともとは物質の変化する現象を研究する学問であった．普通は，物質に熱エネルギーを与えたり，光エネルギーを与えたりすると，変化が起こる．物質 A と物質 B を混ぜたときも，もし，化学反応によって熱が放出されれば，自然に変化が起こる．例えば，塩酸と水酸化ナトリウム水溶液を混ぜれば，発熱しながら，反応が進む．このとき放出される熱エネルギーは**中和熱**と言われる．

$$\mathrm{HCl} + \mathrm{NaOH} \rightarrow \mathrm{NaCl} + \mathrm{H_2O}.$$

ところが，二つの物質が化学反応を起こすこともなく，エネルギーも変化しないにも関わらず，自然に起こる変化がある．例えば，二つの同じ体積の容器を用意して，それぞれの容器に同じ量のヘリウムガスとアルゴンガスを詰めておく（図 2）．もちろん，圧力も体積も温度も同じにしておく．また，外から熱エネルギーが加わらないように断熱材で覆い，外から光エネルギーが加わらないように遮光材で覆ったとしよう．

二つの容器を隔てている境目の壁を取り除くと，どうなるだろうか．もちろん，ヘリウムガスもアルゴンガスも安定な化合物であり，不活性ガスと

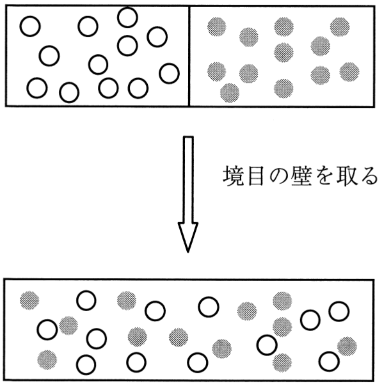


図 2　2 種類の気体は自然に混合する．

言われるくらいだから，化学反応を起こすことはない．また，外からの熱エネルギーも光エネルギーも遮断されているので，気体のエネルギーが増えることもない．全体の体積も変わらないし，圧力も温度も変わらない．それにも関わらず，なんと，自然に，二つの気体が混ざり合うのである．当り前のことと思うかもしれないが，不思議なことはエネルギーを与えなくても混ざり合ってしまうことである．そして，いつまで待っていても，もとの状態，つまり，ヘリウムガスとアルゴンガスがそれぞれの容器に別々に存在する状態にはならないことである．どうしてだろうか？ どうして 2 種類の気体は混ざりたがるのだろうか？ このように，一方向的に変化が起こり，逆の方向に変化が起こらないときに，化学では**不可逆過程**と言う．

3. 状態数とエントロピー

どうして，2 種類の気体が混ざったままで，もとの別々の気体にもどらないかを**確率**で考えてみよう．この問題を分かりやすくするために，それぞれの気体の原子が 4 個ずつ容器に入っているとする．図 3 (a) は境目を取り除く前の状態を表しており，左の容器と右の容器に分かれている．この状態は 1 種類しかないので，状態数を 1 と数えることにする．

もしも，境目の壁を取り除くと，どのようになるのだろうか．つまり，それぞれの気体は自由に左へも右へも移動することができる．境目の壁を取り除いたにも関わらず，もとのように，2 種類の気体が左右に分かれている状態数は 1 である（図 3 (b)）．これに対して，例えば，それぞれの気体の 2 個の原子を交換したときのことを考えてみよう．このときには，それぞれの原子の位置によって，いろいろな状態が考えられる．図には書ききれないけれども，ていねいに数えれば，36 通りの組合せが考えられる．つまり，状態数は 36 になり，境目の壁を取り除いたにも関わらず，壁を取り除く前の状態数の 36 倍である．それぞれの状態になる確率は同じであると考えられるので，ヘリウムガスと

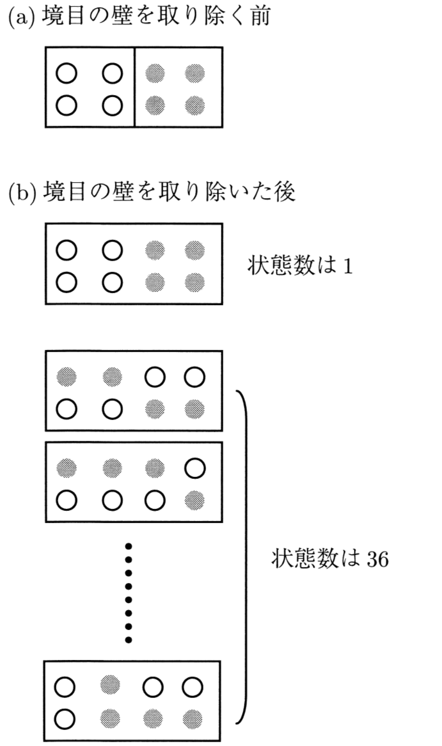


図 3　状態数の違い．

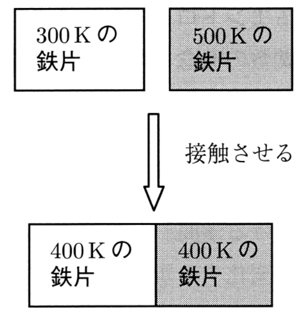


図 4　熱エネルギーの移動．

アルゴンガスがどちらにあるかを観測すると，混合した状態で観測される確率のほうが 36 倍大きい．つまり，境目の壁を取り除いたにも関わらず，もとのように 2 種類の気体が左右に分かれている状態を観測することはほとんどない．したがって，見かけ上，2 種類の気体は混合したように見える．

上記の 2 種類の気体の混合の問題で分かるように，物質の変化を正確に理解するためには，エネルギーだけではなく，状態数，あるいは，確率という物理量を考慮しなければならない．その物理量が**エントロピー**である．エントロピーとは，ギリシャ語で“方向を与える”という意味であり，記号は S で表す．詳しいことは省略するが，状態数を W とすると，エントロピーは状態数の自然対数に比例する．

$$S = k\ln(W). \tag{1}$$

これはボルツマン（L.E. Boltzmann）の考えた式であり，比例定数 k は**ボルツマン定数**と言われる．この式からすぐに分かるように，状態数が多ければ多いほど，エントロピーは大きくなる．

エントロピーを理解するために，もう一つの例をあげよう．同じ物質量で，温度が 500 K の高温の鉄片と，300 K の低温の鉄片を用意して，接触させることにする（図 4）．そうすると，外からエネルギーを加えなくても，自然に，高温の鉄片から低温の鉄片に熱エネルギーが移動して，同じ温度，すなわち，400 K になる．そして，いくら待っていても，もとのように，高温の鉄片と低温の鉄片になることはない．これも，全体のエネルギーは変化しないけれども，エントロピーが変化するために起こった現象である．詳しいことは省略するけれども，温度が T のときのエントロピーは温度の自然対数に比例する．

$$S = C\ln(T). \tag{2}$$

このときの比例定数 C は**熱容量**と呼ばれる定数である．物質の温度を 1 K 上げるために必要な熱エネルギーと考えればよい．

二つの鉄片を接触する前と後の鉄片のエントロピー変化（ΔS）を計算してみよう．高温の鉄片は 500 K が 400 K になったのだから，

$$\Delta S = C\ln(400) - C\ln(500) = C\ln(400/500)$$

となる．一方，低温の鉄片は 300 K が 400 K になったのだから，

$$\Delta S = C\ln(400) - C\ln(300) = C\ln(400/300)$$

となる．したがって，全体のエントロピー変化は，

$$\Delta S = C\ln(400/500) + C\ln(400/300)$$
$$= C\ln(16/15) = C\ln(1.06666) > 0$$

となり，エントロピーの変化量が正の値であることが分かる．どのような温度変化でも，自然対数の括弧の中の数は1以上になるので，エントロピーは常に増大していることが分かる．“宇宙のエネルギーは保存されるけれども，宇宙のエントロピーは増大する”という有名な言葉がある．

4. 物質波の概念

はじめに述べたように，化学は物質の構造や性質を調べる学問である．物質を細かく分析していくと，物質は，やがて，分子や原子にたどりつく．したがって，物質の構造や性質を知るためには，分子や原子の構造や性質を調べればよい．最近，分析技術と量子化学計算は急速に発展し，まるで，目で見るかのごとく，分子や原子などのミクロの世界の解明が可能となっている．

分子や原子の構造や性質を決めているものは電子の存在確率である．電子がどこにどのくらいの確率で存在するかを理解すれば，すべてが理解できると言っても過言ではない．最近の量子化学では，ある程度までの大きさの分子ならば，それが可能となってきた．特に，分子軌道法の理論的な発展によるところが多い．しかし，「電子の存在確率」という概念が提案されたのは，量子論が誕生しようとしていた20世紀半ばのことである．

その当時，光で代表される電磁波は，「実体が波」であるにも関わらず，エネルギーとか運動量という「粒子の性質」をもつことがしだいに明らかになった．古典力学によれば，質量 m をもつ粒子が速度 v で運動しているときには，その運動エネルギーは $(1/2)mv^2$ であり，また，運動量は mv で表される．電磁波は質量 m がないにも関わらず，エネルギーも運動量もあるという結論にいたったのであるから，その当時の物理学者が一斉に驚いたことも理解できる．しかし，もっと，驚いたことに，「実体が波」である電磁波が「粒子の性質」をもつように，「実体が粒子」である電子が「波の性質」をもつと言い出した物理学者がいた．ド・ブロイ（de Broglie）である．彼は「すべての物質は波の性質をもつ」という**物質波**の考え方を提案した．

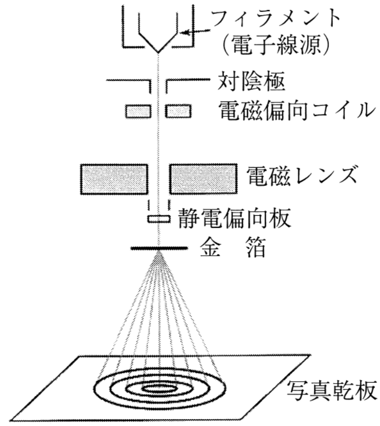


図5　金箔による電子回折．

電子は本当に波の性質をもつのだろうか．これを確かめるための実験がある．どのようにして確かめるかというと，電子を高電圧で加速して，金箔に衝突させる（図5）．そうすると，X線回折の実験と同じように，電子は回折され，そして，重なりあって干渉を起こして，いくつもの同心円ができる．もちろん，X線は電磁波であるから，実体は波である．電子を金箔に衝突させても，X線を金箔に照射させたときと同じように，同心円の回折パターンが得られるということは，実体が粒子である電子が，実体が波であるX線と同じように，波の性質をもっていることの実験的な証明である．

気をつけなければならないことは，「電子の実体が波になったのではない」ということである．電子はあくまでも質量をもった粒子である．その証拠に，電子を1個だけ金箔に衝突させれば，同心円ではなく，1点が観測されるはずである（図6）．電子の実体は波ではなく，あくまでも粒子であり，その存在は点として観測される．それでは，同心円という回折パターンはどうしてできるのだろうか．実は，そこに「確率」の概念が必要となる．

二つ目の電子を金箔に衝突させると，どうなるだろうか？ 古典力学では，ニュートンの運動方程

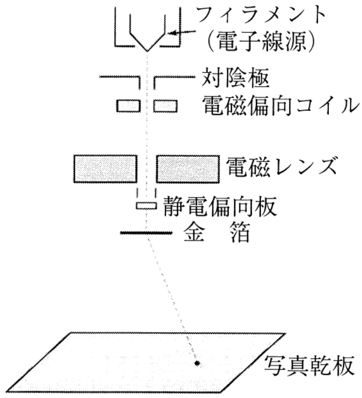


図６　１個の電子の散乱．

式によって，電子の運動に関する解を厳密に求めることができる．ちょうど，太陽と地球と月の位置関係を正確に計算して，日食が何時何分何秒にどこで起こるかを予測できるようなものである．そうすると，何個の電子を金箔に衝突させても，ニュートンの運動方程式に従って，どの電子も同じところに散乱されるはずである（電子は粒子なので，ここでは，回折ではなく散乱という言葉を使っている）．しかし，実際に実験をしてみると，そのようにはならない．電子がどの方向に散乱されるかは全く予測がつかない．

金箔に衝突させる電子の数をしだいに増やしていくと，やがて，同心円のパターンがはっきりと現れるようになる．つまり，一つ一つの電子はどの方向に散乱されるかが決まっていないけれども，どのくらいの電子がどの方向に散乱されるかという**確率**は厳密に決まっているのである．確率が高くなれば，その部分は同心円となって，はっきりと現れる．X 線回折の場合には，X 線の振幅の大きいところが同心円になって現れるから，X 線の振幅が電子の存在確率に対応していると考えるべきである．つまり，「電子の波の振幅は電子の存在確率を表す」と考えてよい．

5. 原子の波動関数

水素原子には１個の電子が含まれている．電子の波の性質とは，どのようなものなのだろうか．詳しいことは省略するけれども，シュレーディンガー方程式 (3) を立てて，それを解くと，電子の波の振幅 Ψ を求めることができる．この振幅のことを**波動関数**と言う．

$$\left[-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)-\frac{e^2}{4\pi\varepsilon_0 r}\right]\Psi = E\Psi. \tag{3}$$

ここで，$\hbar$ はプランク定数を 2π で割った値，m_e は電子の質量，e は電気素量（電子および陽子の電荷の絶対値），ε_0 は真空の誘電率，r は電子の原子核（陽子）からの距離である．

シュレーディンガー方程式は２階の微分方程式なので，得られる波動関数は実関数の場合もあれば，複素関数の場合もある．前述したように，波動関数は波の振幅を表し，電子の存在確率を表すはずであるけれども，波動関数の値が虚数になった場合にはどのように考えたらよいのだろうか．確率が虚数などということは考えにくい．この問題を解決するために，ボルン（M. Born）は「波動関数の２乗は電子の存在確率を表す」という考え方を提案した．つまり，複素関数の場合には，複素共役 Ψ^* をかけた値を存在確率と考える．

$$|\Psi|^2 = \Psi^*\Psi. \tag{4}$$

こうすれば，波動関数の２乗は必ず正の値になるから，存在確率の概念に対応させることができる．

シュレーディンガー方程式 (3) を解いて，水素原子の波動関数を求めるためには，直交座標系 (x, y, z) を極座標系 (r, θ, ϕ) に変換するとよい．詳しいことは省略するが，その結果は，

$$\Psi(r,\theta,\phi) = R(r)_{n,l}Y_{l,m}(\theta,\phi) = L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right)P_l^{|m|}(\cos\theta)\exp(im\phi) \tag{5}$$

となる．ここで，$L_{n+l}^{2l+1}(2r/na_0)$ はラゲールの陪多項式，$P_l^{|m|}(\cos\theta)$ はルジャンドルの陪多項式である．a_0 は**ボーア半径**と呼ばれる定数である．式の中に現れる整数，n, l, m は**量子数**と呼ばれ，いくつかの制限がある．n が１の場合には l も m も０となる．また，n が２の場合には l が１また

表2　水素原子の波動関数.

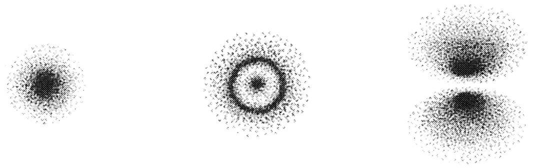
(n,l,m)	名称	波動関数
$(1,0,0)$	1s	$\frac{1}{\sqrt{\pi a_0^3}}\mathrm{e}^{-\frac{r}{a_0}}$
$(2,0,0)$	2s	$\frac{1}{4\sqrt{2\pi a_0^3}}\left(2-\frac{r}{a_0}\right)\mathrm{e}^{-\frac{r}{a_0}}$
$(2,1,0)$	$2\mathrm{p}_z$	$\frac{1}{4\sqrt{2\pi a_0^3}}z\mathrm{e}^{-\frac{r}{a_0}}$
$(2,1,\pm1)$	$2\mathrm{p}_x$	$\frac{1}{4\sqrt{2\pi a_0^3}}x\mathrm{e}^{-\frac{r}{a_0}}$
	$2\mathrm{p}_y$	$\frac{1}{4\sqrt{2\pi a_0^3}}y\mathrm{e}^{-\frac{r}{a_0}}$

は0で，l が1の場合には m は0または ±1，l が0の場合には m も0となる．

それぞれの波動関数を具体的に表 2 に示す．(n,l,m) が $(1,0,0)$ の波動関数は**1s 軌道**と呼ばれ，式から分かるように，r のみの関数であるから，球対称である．(n,l,m) が $(2,0,0)$ の波動関数は**2s 軌道**と呼ばれる．この場合も 1s 軌道と同様に球対称である．ただし，r が $2a_0$ のときには波動関数の値が常に0となる．このように波動関数の値が常に0となる空間を**節**と呼ぶ．ギターなどの弦を押さえて，振動させない点と同じである．(n,l,m) が $(2,1,-1)$, $(2,1,0)$, $(2,1,1)$ の波動関数は**2p 軌道**と呼ばれる．それぞれ，x, y, z が式の中に現れているので，軸対称である．

波動関数を得ることができれば，原子の中で電子がどこにどのくらいの確率で存在するかを知ることができる．しかし，表 2 の式を見ていただけでは，分かりにくい．そこで，ある点における波動関数の2乗の値を図にしてみよう．しかし，$(x,y,z,|\Psi|^2)$ の4次元のグラフを描くことは難しい．そこで，図7のように，xy 平面の断面をとることにする．電子回折の実験で得られた同心円のパターンと同じように，濃いところは電子の存在確率が大きく，薄いところは電子の存在確率が小さいことを意味している．

この図から分かるように，1s 軌道の波動関数は確かに球対称である．中心から離れるにしたがって，電子の存在確率はしだいに小さくなる．2s 軌道の波動関数も，断面が円になっていることから，球対称であることが分かる．ただし，1s 軌道と異なり，中心部分のほかに，$r=3a_0$ のところ（波動関数の2乗の極大値を与える空間）でも，電子の存在確率が高くなっている．一方，2p 軌道の波動関数は球対称ではなく，軸対称である．



1s 軌道　　2s 軌道　　2p 軌道

図7　水素原子の波動関数.

原子と原子が結合して，分子になっても同様である．この場合には，原子の波動関数の線形結合をつくって，分子の波動関数を近似的に求める．これを **LCAO 近似**と言う．Linear Combination of Atomic Orbitals の略である．分子を構成する原子の種類が異なれば，分子の中の電子の存在確率は対称的でなくなる．電子の存在確率が大きいところでは，電気的に負の電荷に偏り，正の電荷をもつイオンなどと化学反応を起こしやすくなる．逆に，電子の存在確率が小さいところでは，電気的に正の電荷に偏り，負の電荷をもつイオンなどと化学反応を起こしやすくなる．どのような化学反応が起こるかを知りたければ，分子の波動関数を計算して，電子の存在確率を調べればよい．

最近，量子化学計算によって，電子の存在確率を求めるコンピュータプログラムソフトが安価で手に入るようになった．物理化学者はもちろんのこと，有機化学者も無機化学者も分析化学者も，ほとんどすべての化学者が量子化学計算を利用している．

参考文献

1) 中田宗隆：『化学—基本の考え方 12 章』，東京化学同人 (1994).
2) 中田宗隆：『量子化学—基本の考え方 16 章』，東京化学同人 (1995).
3) 中田宗隆：『なっとくする量子化学』，講談社 (2001).

（なかた・むねたか，東京農工大学大学院共生科学技術研究院）

分子進化と確率モデル

岸野　洋久

1. 突然変異と淘汰

集団の中は突然変異が蓄積するにつれ多様化していく．そしてあるときは，強い環境の選択圧を受けて集団は細り，これに適応するものが次の世代を形づくっていく．突然変異による多様化と選択圧による絞込みが生物進化の動力源となっている．図 1 は，こうした生物進化を模式的に表現している．

私たちの健康を守り，食料を確保するために，世界中で殺虫剤や抗生物質が散布されている．こうした選択圧に対して，生態系の中に迅速に適応するものが出てきている．ある場合には外からの異物を分解する酵素をコードする遺伝子の制御領域に突然変異が生じ，酵素が異常に生産される個体が出現した[1]．またある場合には，阻害剤に対する抵抗性も出現している[2]．阻害剤は，生命活動の鍵となる生化学反応の担い手になりすましたものである．結合を助ける酵素や相手に突然変異がタンパク質の形を微妙に変化させ，阻害剤よりも役者が効率的に結合できるようになる．

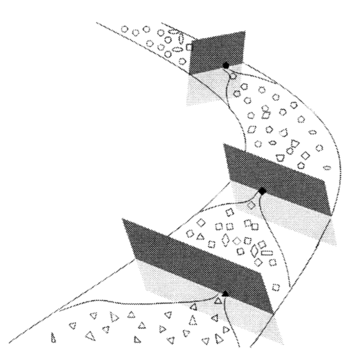


図 1　進化の流れ：突然変異による多様化と環境による選択圧のバランス．

こうした都合のいい突然変異は非常に稀に，確率的に起きるものであるが，選択圧が強い環境下では，それが効率よく選び取られるのである．そこで，生物の適応の背後にある突然変異と集団への定着の過程を追ってみよう．

2. ウィルスをめぐる抗体と宿主細胞の席取り競争

図 2 は SARS ウィルスが抗体および宿主細胞の受容体と結合しているスナップショットを重ね合わせたものである．実際には SARS の結合相手は抗体か宿主細胞のいずれかである．見方を変えると，この両者が SARS をめぐって席取り競争を行っている．

突然変異によるウィルスの適応進化を定量化するために，宿主内のウィルス，抗体，宿主細胞の 3 者系を考える．抗体が結合するとウィルスは消滅し，宿主細胞の受容体と結合するとその細胞はウィルスを自身と同化し，中に取り込む．ウィルスは宿主細胞のゲノムにある転写・発現調節の機構を利用して次世代のウィルスを生成する．ウィルス V，抗体 W，正常細胞 X，感染細胞 Y の宿主内における個体群動態は，簡単な数理モデル

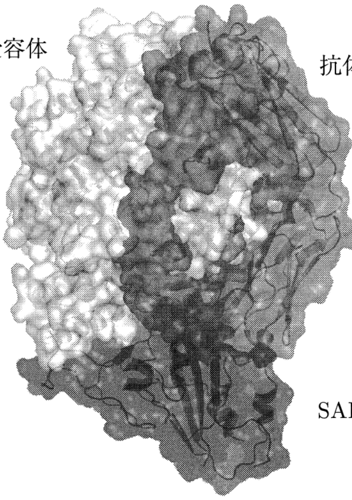


図 2 SARS ウィルスと抗体との結合, SARS ウィルスと宿主細胞の受容体との結合を重ね合わせる. データは PDB (Protein Data Bank) に登録されている.

$$\begin{aligned}\frac{dX}{dt} &= \lambda - dX - \beta XV,\\ \frac{dY}{dt} &= \beta XV - AY,\\ \frac{dV}{dt} &= kY - uV - \beta XV - qVW,\\ \frac{dW}{dt} &= rV - hW - qVW\end{aligned} \tag{1}$$

により表現される.

(1) 式における β は, 宿主細胞と衝突したウィルスが, 細胞に取り込まれる確率と関係している. 同様に, q は抗体と衝突したウィルスが, 系の外に消滅する確率と関係している. それ以外のパラメータは, 3 者系の相互作用とは関係なく働く生成・死滅の機構を表現している. 突然変異により宿主細胞の受容体, あるいは抗体との結合部位における形状や疎水度, 電荷が変化すると, 結合能 β, q が変化する.

数多くのタンパク質やタンパク質の複合体の立体構造が調べられ, データベースに登録されている. このデータベースから, 例えばアラニンやグリシンなどの小さなアミノ酸は構造の安定性を大きく下げることがないため, 比較的多く存在するが, 特にタンパク質の内部に多いこと, これに対して, 大きなメチオニンやリジンなどは表面に位置する傾向があることなどが見えてくる. 結合状態と自由状態でアミノ酸の置かれた局所環境を比較することにより, 結合能を大雑把に見積もることができる. こうして推定された結合能を用いて, 従来型ウィルスの持つ係数から得られるウィルス個体群の大きさと突然変異を受けたウィルスの持つ係数から得られるウィルス個体群の大きさを比較する. これにより, 突然変異の従来型に対する適応度が得られる[3)].

3. 突然変異の運命

集団中に生じた突然変異は, まわりよりも適応度が勝っているものは, 頻度が拡大し, やがて集団に固定される傾向がある. 逆に適応度が劣っているものは, 集団から排除される傾向がある. この傾向の強さは, 適応度の差と集団の大きさで決まってくる.

従来型のアレル 0 (対立遺伝子) に比べ, 突然変異を受けたアレル 1 の適応度を $1+s$ として, アレル 1 の頻度 f の時系列的変化を追ってみる. 遺伝子集団の大きさを N とすると, 突然変異が起きた時点では $f(0)=1/N$ である. t 世代後において $f(t)$ であったとすると, その次の世代では平均的に

$$\begin{aligned}&\mathrm{E}\left[f(t+1) \mid f(t)\right]\\ &= \frac{f(t)(1+s)}{f(t)(1+s)+(1-f(t))}\end{aligned} \tag{2}$$

となる.

集団が有限であることから, 確率性が生まれる. 各遺伝子で世代継承が独立に行われる場合には, 次世代のアレル頻度は二項分布に従う. 図 3 は, 大きさ 10000 の集団中に, 適応度が 2 倍 $(s=1)$ の突然変異が生じた場合のシミュレーション結果を表している. 全体の 80% において新規アレルは集団に固定し, 集団に定着するまでに平均 27.5 世代, 標準偏差 2.3 世代を要している. 他方, 20% の確率で新規アレルは集団から脱落しており, 脱落までの世代数は平均 1.5 世代と短い. マイナー

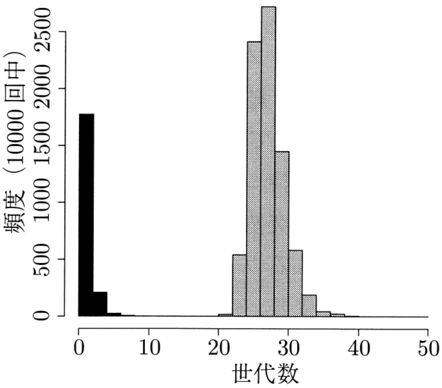


図 3 変異アレルが集団中に固定するまでの世代数（灰色）と集団から排除されるまでの世代数（黒色）の分布．$s = 1$, $N = 10000$.

アレルとしてスタートするために，最初の 1–2 世代がこの突然変異の運命を決定付けていることがわかる．

適応度において大きな違いがない場合には，突然変異の運命が決定するまでに多くの世代がかかる．こうした場合には，世代を連続近似して拡散過程で表現すると見通しがよくなる．(2) 式から移流項を，二項分布の分散により拡散係数を見積もることができる[4]．

4. 宿主と寄生種の軍拡競争と相互情報量

宿主と寄生種の間には絶えず軍拡競争が繰り広げられている．宿主は寄生種の攻撃を回避するよう適応し，寄生種は効率的に寄生するよう適応する．遺伝子配列を種間で比較することにより，適応進化の痕跡を見つけ出すことができる．

バクテリア *Bacillus thuringiensis* の持つ毒素 Bt toxin は，昆虫を死に至らしめる．Cry タンパク質が錐のような役割をなし，中腸の上皮細胞に穴を開ける．現在 100 種類以上の Cry タンパク質が発見されているが，それらは強い宿主特異性を持っている．生態系を大きく歪めることなく，特定の昆虫のみを標的とすることができることから，農薬や遺伝子組換え作物などにも利用されている．

Cry タンパク質は 3 つのドメインからなる構造を持っている（図 4）．ドメイン I は並行する 7 つ

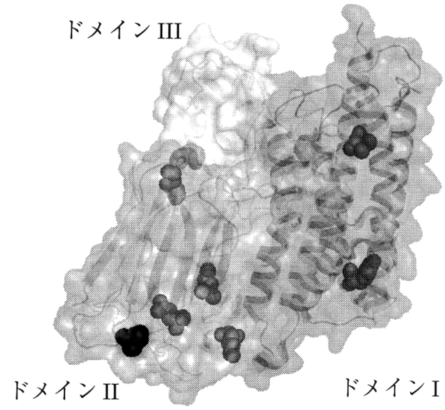


図 4 Cry タンパク質の立体構造．カドヘリンおよびアミノペプチドと共進化したと推測されるアミノ酸残基を原子表示している．灰色は恐らく偽陽性．黒は変異体の実験で毒性との関連が確認されている．

の α ヘリックス[*1)]から構成されており，これが錐の働きをなすと信じられている．ドメイン II, III はそれぞれ 3 つ，2 つの β シートからなるが，宿主の中腸上皮細胞膜にある受容体と結合し，これを梃としてドメイン I が細胞膜に穴を開けると考えられている．特にドメイン II の β シートを繋ぐループに多様性が高く，宿主特異性と関わっていると見なされている．

Cry タンパク質と結合するタンパク質の候補として，カドヘリン様タンパク質（CLP）とアミノペプチド（APN）が実験から示唆されている．ただ，膜上に位置する膜タンパク質であるため，結晶化が難しく，立体構造が得られていない．そこで，共進化したアミノ酸サイトを推定する．宿主と寄生種のタンパク質間の結合部位において，一方のアミノ酸が変異し，構造が微妙に変化すると，これに対抗する形で他方のアミノ酸に変異が起き，再びタンパク質間相互作用に変化をもたらしている．共進化しているサイトの対は，分子系統樹上，同じ枝で変異を起こしている可能性が高いため，2

*1) タンパク質の立体構造を大雑把に表現するために，しばしば 2 次構造を用いる．これは主鎖の繋がりの形状を表現したもので，図中リボンで描かれているように，α ヘリックス，β シート，およびこれらを繋ぐループにより構成される．前者 2 つは構造が堅く保たれるのに対し，ループは柔軟性に富むことが知られている．

つのサイトのアミノ酸のパターンは，独立性から離れてくる．独立性からのずれを相互情報量により定量化する．

配列データベースから，標的昆虫16種のCLPと標的昆虫23種のAPNを抽出し，それぞれこれに対応するCryタンパク質との共進化を調べた．Cry-CLP間で200万のアミノ酸サイト対，Cry-APN間で130万のアミノ酸サイト対の相互情報量を計算したところ，p値が10^{-7}の有意性を持つものがCry-CLP間で100対，Cry-APN間で44対検出された（武内ら，unpublished）．このうち，Cryタンパク質の中で，共進化するアミノ酸が1対1対応して局所的な相互作用が示唆されるものを図4に原子表示している．

灰色のものは実験的な確証が得られておらず，擬陽性である可能性を排除できない．これに対してドメインIIのループ上にある黒く着色したアミノ酸については，同一のループ上のアミノ酸を人工的に変異させたものが毒性を失うことが確認されている．したがって，このループが受容体と結合し，宿主特異性に関わっていた可能性が高いと言える．

5. 分子進化速度

このように，遺伝子配列を比較することにより，適応進化の痕跡を読み取ることができる．特に，第一の取りかかりに，分子進化速度の変化に注目することが多い．遺伝子配列の種間比較から推測される分子進化は，集団に固定した突然変異である．したがって，突然変異率をμ，集団の大きさ（遺伝子数）をN，突然変異の固定確率をfと書くと，分子進化速度rは$r = \mu \times N \times f$である．

有利な突然変異はまれにしか起こらない．いま簡単のために，この確率を無視し，突然変異はまわりとほぼ同等の適応度を持つ中立な突然変異と，まわりに比べ適応度が劣り，集団から排除される突然変異の2種類からなるとする．中立な突然変異の固定確率は$1/N$である．したがって，中立な突然変異の確率をpとすると，分子進化速度は$r = \mu \times p$となる．

世代の長さが変わったり，暴露される放射線の量など環境が変わると，突然変異率μが変化する．例えば，ハエの遺伝子配列を比較すると，ショウジョウバエに近い系統では分子進化速度が速くなっていることがわかる．これはこの系統で世代の長さが短くなったことと関係している．

これに対して，機能的な制約の変化は，中立な突然変異の確率pの変化と関係する．生物が共生するに伴い，機能が分担され，しばしばそれぞれのゲノムの分子進化速度が上がる．遺伝子重複後には2つのコピー遺伝子の冗長性から分子進化速度が上がる．これが遺伝子として用をなさない偽遺伝子となると，速度は高い位置にとどまるが，新しい機能を獲得すると速度が低くなる．

突然変異率の変化は，ゲノム上の全遺伝子が共有するのに対し，機能的な制約の変化は各遺伝子に個別に作用する．したがって，数多くの遺伝子を解析することにより，分子進化速度の変化がどちらの因子によるものか，推測できる．

6. 変動の推定：尤度，階層ベイズ

まずは簡単のため，分子進化速度の推定値の時系列が得られている状況を考えよう．図5はこうした状況をシミュレーションにより表現している．

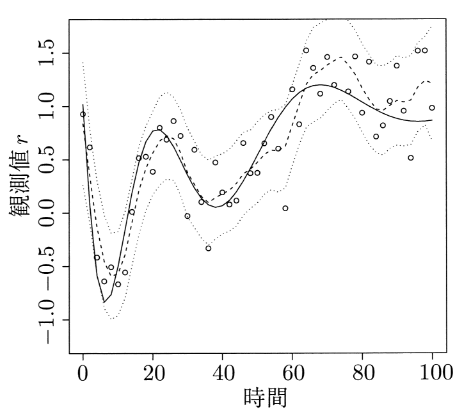


図5 ランダムウォークのベイズ階層モデルによるトレンド推定：実線は真値．破線は事後平均，点線は95% 信用区間を示している．

実線は背後にある変動で，○印が観測値である．滑らかな変動を頑健に推定するために，罰金つき最小二乗法を考えてみる．

観測値を $\boldsymbol{r} = (r_1, \ldots, r_n)$ とし，推定すべきトレンドを $\boldsymbol{T} = (T_1, \ldots, T_n)$ とする．

$$S = \sum_{t=1}^{n} (r_t - T_t)^2 + \lambda \sum_{t=2}^{n} (T_t - T_{t-1})^2 \quad (3)$$

を最小化させることによりトレンドを推定する．第1項はデータへのフィッティングの悪さに対するコスト，第2項は局所変動へのペナルティを表している．λ はペナルティの重みである．これを大きくし過ぎると推定されたトレンドはデータを無視し，ほとんど平らなものとなる．逆に λ を過小に設定すると，トレンドはデータに追随し過ぎ，トレンドとしての意味を持たなくなる．匙加減が大事であるが，種々の λ に対する結果を見て判断するのは，客観性に欠ける．

階層ベイズの枠組みで，このパラメータをデータに基づき評価しつつトレンドを推定することができる．(3) 式には次の統計モデル

$$\begin{aligned}
&\exp\left(-\frac{1}{2\sigma^2}S\right) \\
&\propto \left[\prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(r_t - T_t)^2\right)\right] \times \\
&\quad \left[\prod_{t=2}^{n} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{\lambda}}} \exp\left(-\frac{1}{2\frac{\sigma^2}{\lambda}}(T_t - T_{t-1})^2\right)\right] \\
&\equiv \mathrm{P}\left(\boldsymbol{r} \mid \boldsymbol{T}, \sigma^2\right)\mathrm{P}\left(\boldsymbol{T} \mid \lambda, \sigma^2\right) \qquad (4)
\end{aligned}$$

が対応している．第1項はデータがトレンドのまわりに正規分布するとする尤度である．尤度はデータを生成する確率構造をモデル表現したもので，その対数をとった対数尤度は真の分布とモデル分布の乖離度を表現する Kullback-Leibler 情報量と関係している[5,6]．第2項はパラメータであるトレンドがランダムウォークするとする事前分布である．λ はランダムウォークの単位時間あたり分散を規定している．事前分布を規定するパラメータは超パラメータと呼ばれる．超パラメータに分布を導入することにより，事前分布に対する不確実性を許し，頑健な事後分布を得ることができる（図 5）．一般には分子進化速度は直接観察されず，遺伝子配列データから推定する．(4) 式の最右辺の第1項を遺伝子配列の尤度に置き換えることにより，遺伝子配列から分子進化速度の変動を推定する階層ベイズモデルを構築することができる[7]．

7. 分子進化の統計モデル

進化の過程で，遺伝子配列のそれぞれのサイトのアミノ酸が確率的に変化していく．この変化はマルコフ過程によりモデル化することができる．それぞれのサイトの分子進化が独立だとすると，長さ n の s 本の配列データ

$$\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_h, \ldots, \boldsymbol{X}_n) = \begin{bmatrix} X_{11} & \cdots & X_{1h} & \cdots & X_{1n} \\ \vdots & & \vdots & & \vdots \\ X_{p1} & \cdots & X_{ph} & \cdots & X_{pn} \\ \vdots & & \vdots & & \vdots \\ X_{s1} & \cdots & X_{sh} & \cdots & X_{sn} \end{bmatrix}$$

の対数尤度は，それぞれのサイトの対数尤度の和として得られる．種分化の後，生殖的に隔離された両系統の遺伝子は独立に進化すると仮定すると，各サイトの尤度は

$$\begin{aligned}
&\mathrm{P}\left(\boldsymbol{X}_h \mid \boldsymbol{b}, \boldsymbol{\theta}\right) \\
&= \sum_{Z_{i_0 h}} \pi_{Z_{i_0 h}} \prod_{j \in node(\mathcal{T}) \setminus i_0} \sum_{Z_{jh}} P_{Z_{anc(j)h}, Z_{jh}}(b_j)
\end{aligned}$$

と，平衡確率と隣接節の状態の推移確率の積で表現される．内部節の状態が未知であることから，各状態の確率を足し合わせる．推定すべきパラメータは各枝の期待置換数 $\boldsymbol{b}$，分子進化速度行列を規定するパラメータ $\boldsymbol{\theta}$ である．$node(\mathcal{T})$ は系統樹 $\mathcal{T}$ の節を表し，i_0 はその根である[8,9]．

図 6 は，鱗翅目（チョウとガ）における不飽和化酵素 desaturase の分子系統樹を示している．この酵素は脂肪酸（C_mH_nCOOH）の炭素結合の一部を二重結合にし，不飽和脂肪酸を生成する．選択する基質脂肪酸の長さと不飽和化する位置に

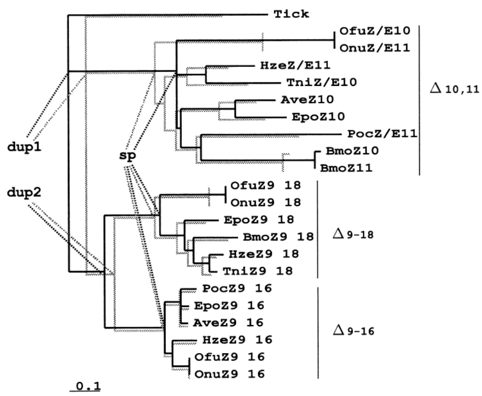


図 6　鱗翅目における desaturase の分子系統樹．dup1, dup2 は 2 回の遺伝子重複，sp は解析した昆虫の共通祖先を示しており，3 つの sp は同時点である．枝の長さはサイトあたりのアミノ酸置換数を表す．黒線は進化速度がサイト間で不均質とする分析，灰色はサイト間で一定であることを仮定した分析．

より分類される．一般に不飽和脂肪酸は飽和脂肪酸よりも融点が低いことから，耐冷性と関係し，$\Delta 9-18$*2) は鱗翅目を越えた昆虫にも共有されている．これに対し，基質の短い $\Delta 9-16$ は沸点も低く，揮発性のある性フェロモンを合成する．鱗翅目昆虫のみに存在することから，$\Delta 9-18$ の遺伝子重複により $\Delta 9-16$ が誕生したと考えられている．

2 回の遺伝子重複が起きており，各生物は 3 種類の似た配列を持っている．$\Delta 10, 11$ の系統は枝が長く，分子進化速度が速い．これは基質の部位が定まらず，機能的な制約が緩んでいることが窺われる．sp で示されている節は，同一の種分化に対応しており，時点は同じである．2 回目の遺伝子重複後，sp 以降においては，$\Delta 9-16$ で分子進化速度が落ちていることが読み取れる．このことは，重要な機能を獲得したことを裏付けている．

タンパク質の表面は変化の自由度が高い．これに対して，タンパク質の内部は多くのアミノ酸と相互作用しており，構造上の安定性を保持しつつ変化する必要がある．このため，速度が遅い．分子進化速度が配列内で均質であることを仮定した推定（灰色）は最大対数尤度が -6125 であるのに対して，配列内の不均質性を考慮に入れた推定（黒）の最大対数尤度は -5861 と，適合度が大幅に改善された．枝の長さが分子進化のモデルに大きく影響されることから，適合度の高いモデルを用いることが大切である．立体構造の情報が得られる場合には，この情報を階層ベイズモデルに組み込むことによりタンパク質のどの領域に強い選択圧がかかったか，直接的にパラメトライゼーションできる．これにより擬陽性を低く抑えつつ検出力を高めることができよう．

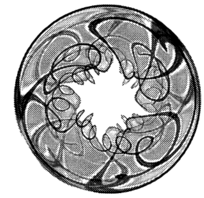
*2)　基質の長さが 18 の脂肪酸の 9 番目の炭素を二重結合にする酵素．

参考文献

1) Daborn et al. (2002). A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*. **297**: 2253–2256.
2) Funke, T., Han, H., Healy-Fried, M.L., Fischer, M., and Ernst Schönbrunn, E. (2006). Molecular basis for the herbicide resistance of Roundup Ready crops. *Proc. Nat. Acad. Sci. USA* **103**: 13010–13015.
3) Watabe, T. and Kishino, H. (2010). Structural considerations in the fitness landscape of a virus. *Mol. Biol. Evol.* in press.
4) Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics*. **47**: 713–719.
5) Kullback, S. (1959). "Information Theory and Statistics". Wiley, New York.
6) Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* **AC-19**: 716–723.
7) Thorne, J.L., Kishino, H. and Painter, I.S. (1998). Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**: 1647–1657.
8) Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**: 368–376.
9) Yang, Z. (2006). "Computational Molecular Evolution". Oxford University Press, Oxford. 藤博幸，加藤和貴，大安裕美訳『分子系統学への統計的アプローチ』共立出版．

（きしの・ひろひさ，東京大学大学院農学生命科学研究科）

沢山ある「コト」に確率で迫る



樺島 祥介

1. はじめに

独立研究科に勤務しているため，受験生を募るための説明会を数回開催することが春先の恒例行事となっている．そこでは研究室で行っている研究をポスターに仕立てて教員自ら説明をする．受験生集めが目的なので参加者は学部4年生がほとんどである．彼らに対して「うちでは確率を使ったうまい情報処理の方法を研究している」といった類の話をするのだが，情報工学科の学生を含めて多くの学生には確率と情報がなぜつながるのかすぐにはイメージできないようである．

もちろん，情報の数理を研究している研究者にとって確率と情報が切っても切れない関係にあることは常識である．だからといって，最近の若い奴らはものを知らなくて困る，と嘆くのも当の学生には気の毒な話であろう．世の中で情報技術といえばケータイであり，ネットであり，ロボットである．なるほど，外側を見る限り確率との関わりなど目につかない．そういったインフラ系の情報技術に関しては確かに確率なんて知らなくても多くの場合事が済んでしまうのである．

実際のところ，確率と情報をめぐるこうした認識は，程度の違いはあれ，つい最近まで専門家の間でも似たようなものであったのではないだろうか．確率と情報は確かにその基礎付けや数理的側面では密接に関係しているものの，技術としては使えない．こうした見方は教員サイドでも半ば常識化しており，このことが確率に関わる講義が概して重視されない国内の情報工学教育に反映されていると思われるからである．

ところが1990年代の半ばに生じた情報符号化に関する技術革新を契機に少なくとも縁の下の力持ち的な基盤技術については確率が最先端の使える技術として利用され始めている．本稿では，こうした確率と情報をめぐる一つの紆余曲折についてその大筋をざっくりと振り返りたい．話の鍵を握るのは“沢山ある「コト」”である．

2. 確率と予測・推論

2.1 最適な予測・推論：曲がったコインの問題

簡単な問題から始めよう．確率 p で表が出ることがわかっている曲がったコインがある．このコインを投げたときに出る目を当てる．どのように予想すればよいだろうか？

答えは至極真っ当である．$p > 1/2$ なら「表」，$p < 1/2$ なら「裏」，$p = 1/2$ なら好きなように「表」または「裏」を予想すればよい．では，なぜ，これがよい予測なのだろうか？ 直感的に明らかすぎるので，こうしたことの「証明」は何を示せばよいのか，また，どこから手をつければよいのか，かえって見当をつけにくくなるなることがある．確率を使うとこうした問題にすっきりとした解答を与えることができる．

まず，予測の方法として確率 $(0 \leq)q(\leq 1)$ で

「表」と答える混合戦略を考える．二者択一の予測を行う際，戦略の良し悪しを評価する尺度は予測が当たる確率であろう．表が確率 p で出ることが仮定されているのでこれは表と答える確率 q の関数として

$$
\begin{aligned}
L(q|p) &= pq + (1-p)(1-q) \\
&= (2p-1)q + 1 - p
\end{aligned}
\tag{1}
$$

により与えられる．

最適な戦略はこれを最大化することで求まる．式 (1) は q の 1 次の係数 $2p-1$ の正負に応じて最大値を与える点 q^* の位置が変化する．$p > 1/2$ では $q^* = 1$，$p < 1/2$ では $q^* = 0$，$p = 1/2$ では q^* は $0 < q < 1$ の任意の値となる．先に述べた予測法はこの結果からすぐに導かれる．

2.2 経験と知識をつなげる：ベイズの公式

簡単な例であるが，上に述べた問題は不確実性を含む状況において知識を予測や推論と結びつけるためのひな形を与えている．対象に関する知識を確率によって表現する．これを使って戦略の良し悪しを期待値として評価する．後はその期待値を最適にする戦略に基づいて予測や推論を行えばよい．

つまり，対象を表現する確率が得られてしまえば，後はすべて同じ方針で問題解決をはかることができる．しかしながら，現実の問題では，表が出る確率は70%である，とわかっている状況などほぼあり得ない．そこで，経験や観測によって得られたデータを確率で表現される知識に結びつけることが重要になる．

こうした作業で重要な役割を果たすのが**ベイズの公式**

$$
P(B|A) = \frac{P(A|B)P(B)}{\sum_B P(A|B)P(B)} \tag{2}
$$

である．ここで，A は経験や観測により得られたデータ，B は未観測のデータを表す確率変数である．$P(A|B)$ は B が与えられた際に A が生じる条件付き確率を表し，これは事前の実験や考察により与えられているとする．$P(B)$ はデータ A を観測する前の B に関する知識を表す確率であり**事前確率**と呼ばれる．一方，$P(B|A)$ はデータ A を得ることにより変化した B に関する知識を表す確率であり**事後確率**と呼ばれる．式 (2) は確率的な知識が経験によってどのように更新されるのか，を表す一般的な公式を与えている．

コイン投げ問題に当てはめてみる．表が出る確率 p に関する知識を得るために実際に n 回コインを投げたところ k 回表が出た状況を想定しよう．まず，初めに p に関する知識が全くないとすれば一様分布 $P(p) = 1$ $(0 \le p \le 1)$ を想定するのが自然であろう．一方，コイン投げの結果が統計的に独立であるとすれば p が与えられたとき n 回のうち k 回表が出る条件付き確率は $P(k;n|p) = n!/(k!(n-k)!)p^k(1-p)^{n-k}$ である．これらの結果をベイズの公式に当てはめ整理するとデータを得た後の p に関する知識が事後確率

$$
P(p|k;n) = \frac{p^k(1-p)^{n-k}}{\int_0^1 dpp^k(1-p)^{n-k}} \tag{3}
$$

によって与えられることがわかる．

この確率的に表現された知識の活かし方にはいくつかの方法がある．ここでは式 (3) を経由することで経験によって得られたデータと予測を直接つなげることを考えよう．そのためには実際に出るコインの目だけでなく，その背後にある p 自体も確率変数と考え，$L(q|p)$ ではなく，事後確率 (3) を用いて期待値評価した

$$
\begin{aligned}
L(q|k;n) &= \int dpL(q|p)P(p|k;n) \\
&= \frac{2k-n}{n+2}q + \frac{n+1-k}{n+2}
\end{aligned}
\tag{4}
$$

を最大化するように予測を行えばよい．このような予測，もっと一般に，行動決定の方法は**ベイズ決定方式**と呼ばれている．

ベイズ決定方式は事前確率 $P(B)$，条件付き確率 $P(A|B)$ が正しくモデル化されているという前提の下，戦略の良し悪しを表すコスト関数の期待値を最適にする意味で最適な行動決定法を与えている．

3. 沢山あるコトと予測・推論

3.1 沢山あるコトはイケテイル：シャノン理論

コイン投げ問題は簡単で理解を助ける例題としては適当である．しかしながら，これだけではベイズ決定方式がどれほど優れた方法なのかイマイチよくわからないというのが大方の見方であろう．こうした見方に対して，確率に基づく行動決定法の予想を超えた能力の高さを最初に示したのはおそらくシャノンによる**情報理論**である．そこでは沢山のコト（=情報）を扱うことが能力の向上に本質的な役割を果たしている．

コイン投げ問題の延長にある話題として**通信路符号化定理**を取り上げることにしよう．**通信路符号化**とはノイズのある通信路を使って情報伝達を行わなければならない状況において，冗長な情報表現に変換することでノイズに対する耐性機能を持たせる符号化方法のことである．

簡単のため，送信者が単一の通信路を使って 2^K 種類のラベル $m=1,2,\ldots,2^K$ を受信者に伝達する状況を考える（図1）．以下，これを**メッセージ**と呼ぶ．ただし，通常の通信路ではメッセージをそのまま送ることはできない．ここでは，理論面でも応用面でも重要なビット値 $x \in \{0,1\}$ を入力とする2元通信路を想定する．また，問題を単純化するため x を入力した際には，その前後の入出力の結果によらず独立に条件付き確率 $P(y|x)$ にしたがって出力信号 $y \in \mathcal{Y}$ が出力される定常無記憶性を仮定する．出力値の値域 $\mathcal{Y}$ は $\{0,1\}$ でなくてもかまわない．

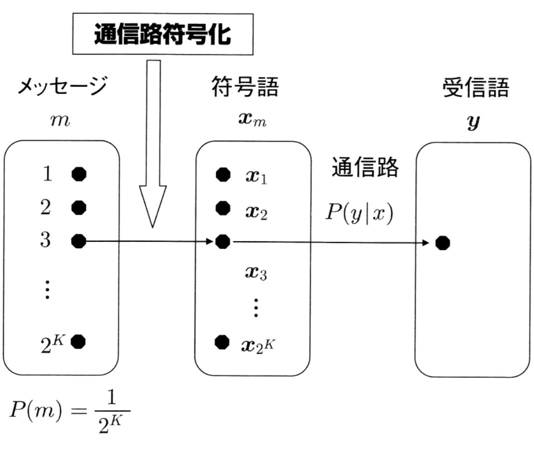


図1　通信路符号化の概念図．

この通信路を使って情報伝達を行うためにメッセージ m のそれぞれをビット値からなる N 次元ベクトル（ビット列）$\boldsymbol{x}_m=(x_{1m},x_{2m},\ldots,x_{Nm})$ $(x_{im}\in\{0,1\}; i=1,2,\ldots,N, m=1,2,\ldots,2^K)$ に変換する．この変換規則を与えるのが符号化であり，しばしば，$\boldsymbol{x}_m$ を**符号語**，対応を表す符号語の組 $\mathcal{C}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{2^K}\}$ を**コードブック**，あるいは，単に**符号**と呼ぶ．2^K 種類のラベルをベクトルで表現するには長さが K だけあれば十分である．ただし，通信路符号化ではより長い $N \geq K$ の長さを持つ符号語を用いて冗長に表現する．比 $R=K/N$ は**レート**と呼ばれ，符号語1ビットあたりに含まれているメッセージに関する情報の密度（あるいは，冗長性の逆数）を表している．受信者は受信した N 次元ベクトル（受信語）$\boldsymbol{y}=(y_1,y_2,\ldots,y_N)$ $(y_i\in\mathcal{Y}; i=1,2,\ldots,N)$ に基づいて送信者が送ったメッセージ m を推定する．この作業は**復号**と呼ばれる．復号結果の良し悪しは符号化方法，つまり，符号 $\mathcal{C}$ の作り方に依存する．通信路符号化の研究では信頼性の高い通信を実現するため $\boldsymbol{y}$ に基づく m の推定精度がなるべく高くなるような符号の設計が求められる．

符号 $\mathcal{C}$ が固定されている場合，最適な復号は以下のようにベイズ決定方式によって実現される．送信したメッセージが m であったとする．これに対し，確率 $q(m')$ で m' を復号結果とする混合戦略を考える．最大化すべき目的関数は復号が成功する確率である．正解 m は確率 $q(m)$ で生成されるので，これは

$$L(q(\cdot)|m)=q(m) \tag{5}$$

と表現される．一方，ベイズの公式を用いると $\boldsymbol{y}$ を受信した後のメッセージ m の事後確率は

$$P(m|\boldsymbol{y})=\frac{P(m)\prod_{i=1}^{N}P(y_i|\boldsymbol{x}_{im})}{\sum_{m'}P(m')\prod_{i=1}^{N}P(y_i|\boldsymbol{x}_{im'})} \tag{6}$$

と評価される．式(6)に関して式(5)を平均する

と $\boldsymbol{y}$ を受信した下での目的関数の期待値

$$L(q(\cdot)|\boldsymbol{y}) = \sum_m q(m)P(m|\boldsymbol{y}) \tag{7}$$

が得られる．$\boldsymbol{y}$ を受信した下での最適な復号戦略はこれを最大化することで求まる．式 (7) は $P(m|\boldsymbol{y})$ を確率分布 $q(m)$ で平均したものである．$P(m|\boldsymbol{y})$ を最大にする m を $\widehat{m}(\boldsymbol{y})$ と記すことにしよう．「平均値 $\leq$ 最大値」の関係を利用すると式 (7) が最大化されるのは与えられた $\boldsymbol{y}$ に対して $q(m)=1$ $(m=\widehat{m}(\boldsymbol{y})), 0$ $(m \neq \widehat{m}(\boldsymbol{y}))$ とする場合であることが導かれる．つまり，事後確率の最大化を行えばよい．これはしばしば**事後確率最大化復号**[*1)]と呼ばれる．

ここまでの議論はコイン投げ問題と同様，確率を用いれば最もよい推論を行うことができるという話である．では，どのくらいよいのか？ その限界はどこにあるのか？ この問いに答えるのは容易ではない．なぜなら，一般に $\widehat{m}(\boldsymbol{y})$ は $\boldsymbol{y}$ に複雑に依存しておりその振舞いの解析は一筋縄ではいかないからである．シャノンの理論がスゴイのは，この種の問いに初めて明確な解答を与えたところにある．

具体的には，最良の符号に対して復号が失敗する確率の上界を評価する．その際，「最小値 $\leq$ 平均値」の関係に着目し，最良符号の失敗確率をランダムに構成した符号に関する復号失敗確率の平均値で上から押さえる，というアイデアを用いる．情報理論における別の考察によりあらかじめ $P(m)=2^{-K}$ となるように情報表現を工夫しておくと情報伝送効率が最適になることが知られているのでこれを仮定する．かなり込み入った議論を必要とするので詳細には立ち入らないが，通信路符号化定理では，最良符号の失敗確率 P_{e} に対して一般に不等式

$$P_{\mathrm{e}} \leq 2^{-NE(R)} \tag{8}$$

の成立が示される．上界を特徴付ける指数 $E(R)$ は**信頼性関数**と呼ばれ通信路の特性を表す条件付き確率 $P(y|x)$ の汎関数として与えられる．信頼性関数は一般に

$$E(R)\begin{cases} =0, & R>C \\ >0, & R<C \end{cases} \tag{9}$$

のように振る舞う．C は**通信路容量**と呼ばれ通信路ごとに定まる情報伝達効率の限界値を表す．

この結果は，$R<C$ であれば，レート $R=K/N$ を一定に保ちながら符号長 N を十分大きくしていくことで，原理的には，復号が失敗する確率を任意に小さくできることを意味している．受信する信号一つ一つは確率的なゆらぎを含んでいるのにも関わらず沢山をまとめて処理すれば必要最小限の長さよりも高々数倍程度長い符号に表現を変換するだけでほぼ確実に正確な情報伝達を行うことができるのである．

量は質を変える，というフレーズをまさに地で行くこの結果は情報処理における“沢山あるコト”の有用性を示した最初の成果であった．その後，1940 年代後半から 50 年代に掛けて，確率的な枠組みに沢山あるコトの視点を導入することにより様々な情報処理の形態に関してその可能性，原理的限界が次々と明らかにされていった．

3.2 沢山あるコトはツカエナイ：次元の呪い

シャノンの理論は，確率をうまく利用すれば理論的に最適な情報処理が実現できることを示唆している．しかしながら，冒頭で述べたとおり，この理論は情報量概念の基礎付けや情報処理の枠組み作りには利用できるものの，実際的な技術の開発にはあまり役に立たないと初期の段階で判断されてしまった．その理由は，やはり“沢山あるコト”に関係している．

通信路符号化定理で情報伝達の正確さを担保しているのは符号長 N が十分大きいという前提である．ところで，ランダムに構成した符号を用いる際のコストを考えてみよう．ランダムに構成するのだから，符号を表現するにはコードブック $\mathcal{C}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{2^K}\}$ を丸覚えしておくほか方法はない．ランダムな N ビットのベクトルを 2^K 個記憶する必要があるのでこれにかかる記憶量は

*1）一様な事前分布 $P(m)=2^{-K}$ の場合は尤度の最大化を意味する．その場合は最尤復号と一致する．

$O(N\times 2^K)$ 程度である．また復号を行う際には，事後確率 $P(m|\boldsymbol{y})$ を最大にするメッセージを探索する必要がある．ランダムな符号語を用いるとすべての符号語を枚挙する以外方法はないのでこれには $O(2^K)$ 程度の計算コストが掛かる．失敗確率を減らしたければこれを十分大きな $K=NR$ に対して行う必要がある．$K=100$ とした場合，必要な記憶コスト，計算コストはアボガドロ数 6×10^{23} をも軽く超えてしまう！ これではとても実用にはツカエナイ．

ベクトルの長さや次元の増加に伴い，処理に必要なコストが指数関数的に増大し実際のシステムでの利用が困難になる問題を情報工学では一般に**次元の呪い**と呼ぶ．確率に基づく予測や推論は，ほとんどの場合，"次元に呪われた" 厄介な問題なのである．通信路符号化の問題は "沢山あるコト" の有用性と同時にその利用の難しさを強く印象付けた最初の例であったとも言える．

こうした事情のため，実用的な符号の研究は許容できる計算コストが制約となり，1960 年代以降，短い長さでできるだけ性能の高い符号の設計に主眼がおかれることとなった．そこでは長さ N，レート R の符号を $(1-R)N$ 個の線形制約式

$$\sum_{i=1}^{N} A_{\mu i}x_i = 0 \ (\mathrm{mod}\ 2) \tag{10}$$

$(\mu=1,2,\ldots,(1-R)N)$ を満たす，0 または 1 を要素とする長さ N のベクトル $\boldsymbol{x}=(x_i)$ の集合として特徴付ける．ただし，mod 2 は 2 を法とする演算である．$A_{\mu i}$ は 0 または 1 の値をとり，それらを要素とする $(1-R)N\times N$ 行列 $\boldsymbol{A}=(A_{\mu i})$ は**パリティ検査行列**と呼ばれる．

符号の性能はパリティ検査行列 $\boldsymbol{A}$ の作り方によって定まる．1960 年代初頭，その設計にそれまで実用とは無縁と考えられてきたガロアによる有限体の理論が有用であることが示された．その結果，実用的な符号の研究ではシャノン流の確率に基づく考察は影を潜め，代わりに，ガロア理論を駆使する代数的符号理論が勃興することとなった．この状況は 1990 年代の初めまで続くことになる．

4. それでも沢山あるコトはイケテイル

ところが，1990 年代の半ば，符号の研究は再び大きな転換期を迎える．ここでも "沢山あるコト" が鍵となる役割を果たしている．

この転換期では，当初，実験による試行錯誤が先行していたのだが，現在の視点から見直してみると，そこで提案された手法はパリティ検査行列に基づく線形符号とベイズ決定方式による復号を組み合わせたものと解釈することができる．符号の研究に確率の視点が戻ってきたのである．

低密度パリティ検査符号と総称されるこの「新しい符号」の研究では各符号語はパリティ検査行列の制約式を満たすベクトルの一様分布から生成されると考える．これがベイズの公式における事前確率を与える．受信語 $\boldsymbol{y}=(y_i)$ を受け取った後は，式 (6) にならい，この事前確率に尤度 $\prod_{i=1}^{N} P(y_i|x_i)$ を掛け更新することで符号語 $\boldsymbol{x}$ の事後確率を求める．問題は少ない計算量でこれからどうやって復号するかである．

この解決には確率の計算に関する発想のコペルニクス的転回があった．変数が沢山ある場合に確率の計算が難しくなるのは，一般に変数が相互に依存しているため，各変数について最適化する，あるいは，和をとる，といった操作が他の変数に次々と影響を及ぼし，結果として局所的な手続きを組み合わせるだけでは手に負えなくなってしまうからである．低密度パリティ検査符号における変数間の依存関係は制約式 (10) によって決まる．この関係は一般に**2 部グラフ**（2 種類のノード（節）をエッジ（辺）で結ぶことでできるグラフ）によって表現できる（図 2 (A)）．

ところで，確率の計算は "一般に" 難しいが，計算が容易になる "特別な" 場合がある．それは 2 部グラフの表現が**循環経路**（**ループ**）を含まないツリー（図 2 (B)）となる場合である．そのような場合には**確率**（**あるいは信念**）**伝搬法**と呼ばれるアルゴリズムを用いるとほぼグラフの大きさに比例する程度の計算量で多変数の確率分布に関する最大化や期待値の評価が可能なことが 1980 年代

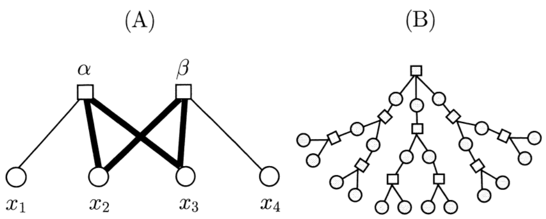


図 2 (A)：式 (10) を表現する 2 部グラフの一例．□は制約条件を，○は符号語の要素をそれぞれ表す．この例では α: $x_1+x_2+x_3=0$, β: $x_2+x_3+x_4=0$（いずれも 2 を法とする演算）を意味している．太線はループを表す．(B)：ツリー（ループを含まない 2 部グラフ）の一例．

の研究によって示されている．面白いことにこのアルゴリズムはループがないグラフに対して少ない計算量で厳密な結果を導くことが保証されているのと同時に，ノード間で計算を局所的に伝搬させる方法であるためグラフ内にループがある・ないといった大域的な性質とは無関係に任意の確率分布に対して適用することができる．ただし，一般のグラフに適用した場合には近似計算にすぎず，どの程度信頼できる結果が得られるか，については別途考察する必要がある．

ここで各行・各列あたりの 1 の個数を定めパリティ検査行列 $\boldsymbol{A}$ をランダムに生成することを考える．グラフの言葉で表現するとこれは 2 種類のノードに対し，1 ノードあたりに結合するエッジの数を定めランダムに 2 部グラフを生成することに対応する．ネットワーク科学の知見によればランダムに生成された 2 部グラフ内の典型的なループの長さはグラフ内のノード数 N に関して $O(\ln N)$ で増大する．つまり，符号の長さ N を十分大きくした符号では典型的なループは長くツリーでよく"近似できる"ことが期待される．このような考察をふまえて，符号長を従来主流であった代数的符号と比較すると桁外れに長く（$N=10^6$ など）し，ランダムに生成した低密度パリティ検査符号に確率伝搬法による近似的復号を組み合わせる．すると，通信路容量に肉薄する性能が得られることが大規模な実験により実証されたのである．現在では，デジタルテレビの衛生通信，大容量ハードディスク等の標準規格となるなど，「低密度パリティ検査符号＋確率伝搬法による近似的復号」は最先端の実用的技術としてすでに不動の地位を築いている．

「計算」という観点から考えると，沢山あるコトに確率で迫るのは難しい，というのが従来の情報工学の常識であった．それが"沢山"を逆手に取ることによって解決される．低密度パリティ検査符号の登場は符号研究のみならず情報工学分野全般の常識をくつがえす画期的な出来事であった．

5. おわりに

確率に基づく予測や推論と"沢山あるコト"との関わりについて述べた．確率の基本公式であるベイズの公式は情報工学において経験と知識をつなげる汎用的な枠組みを与える．シャノンによる情報理論はこうした確率的な予測や推論の枠組みでは大規模化することが性能向上の鍵となることを示唆している．残念ながら，確率の大規模化は一般に計算の実行を難しくする．ところが，最近の符号研究の進展は沢山あるコトで逆に計算困難の問題を実際的に解決できる場合があることを示している．そういった趣旨の話を紹介した．

ところで，前節で低密度パリティ検査符号を「新しい符号」とカッコつきで表現した．これには理由がある．正確を期すと，この符号は 1960 年代初頭，代数的符号の登場と同時期にギャラガーによって提案された「古い符号」だからである．当時，ギャラガーは現在の確率伝搬法のプロトタイプとも言える近似的復号法まで提案しているのだが，当時の計算資源ではその潜在的能力を十分に引き出すことができず，彼の提案はその後 30 年以上ほぼ忘れられた存在となっていたのである．

近似的に計算すればイイじゃない，という素朴な発想にもう一度たどりつくまでに 30 年である．この事実の意味するところはかなりフカイと思うのだが，いかがだろうか．

（かばしま・よしゆき，東京工業大学大学院総合理工学研究科）

学習と確率
記述長最小原理の視点から

山西　健司

1. 学習，確率，符号

財務データ，アクセスログ，トランザクション，センサーデータ…．巷にはデータがあふれている．一見ただの数値列，もしくは文字列にすぎないデータにもそこには「構造」が潜んでいる．例えば，データの出方にどんな特徴があるか？ データが何個のクラスにまとまって現れるか？ というような構造である．本稿では，与えられたデータから，そこに潜む内在的構造を抽出することを問題にする．このような営みを**学習**と呼ぶ（図 1）．本稿では，こうした学習に記述長最小原理の考え方から迫ってみたい．

学習にはモデルが必要である．モデルとはデータの内在的構造を表現する，ある特定の表現形のことである．そのモデルとしては，**確率モデル**を考えることが自然である．

確率モデルとはデータ生成に関する確率分布のことである．$\mathcal{X}$ をデータの定義域とし（仮に有限とする），X を $\mathcal{X}$ に値をとる確率変数とすると，確率モデルは確率分布 $P(X)$ のように表される．以下では，確率変数を表すときには大文字 X で，実現値を表すときには小文字 x で区別する．

今，与えられた観測データ列（仮に n 個とする）$x^n = x_1 \cdots x_n$ は，各 x_i がそれぞれ独立に確率分布

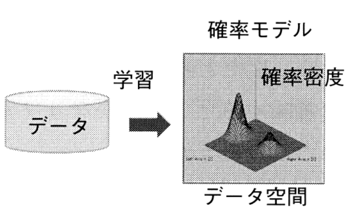


図 1　確率モデルの学習．

$P^*(X)$ に従って発生しているとする（P^* を**真の分布**と呼ぶ）．そこで，真の分布を知らずにデータの生成分布を $P(X)$ と学習したとしよう．このとき，x^n に対する $P(X)$ の良し悪しを

$$-\log P(x^n) = -\log \prod_{i=1}^{n} P(x_i)$$

で測るものとする．ただし，対数の底は 2 であるとする．これは P の x^n に関する**対数損失**と呼ばれる量である．この量が大きければ大きいほど，学習が良くないと考えるわけである．

この量を $P^*(X^n)$ に関して平均した量である**平均対数損失**を調べよう．長さ n のデータ列の P^* に関する平均を $E^n_{P^*}$ と表すと，これは次式のように表される．

$$E^n_{P^*}[-\log P(X^n)] \;=\; H_n(P^*) + D_n(P^*||P).$$

ここに，

$$H_n(P^*) = E^n_{P^*}[-\log P^*(X^n)],$$
$$D_n(P^*||P) = E^n_{P^*}\left[\log \frac{P^*(X^n)}{P(X^n)}\right]$$

である．$H_n(P^*)$ は P^* に関するエントロピー，$D_n(P^*||P)$ は P^* と P の Kullback-Leibler の距離と呼ばれる量である．ここで，Kullback-Leibler の距離に関する一般的な性質として，$D_n(P^*||P) \geq 0$ であり，等号は $P = P^*$ のときに限り成立する．これにより，

$$E^n_{P^*}[-\log P(X^n)] \;\geq\; H_n(P^*) \qquad (1)$$

であり，等号は $P = P^*$ のときに限って成立し，最小値はエントロピーに等しくなることがわかる．

ところで，対数損失 $-\log P(x^n)$ という量は情報理論的に別の意味をもっている．それは x^n の**符号長**と

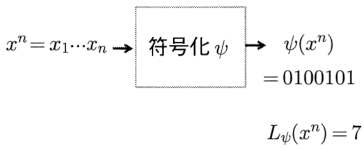


図2　符号化.

いう意味である．ここで，符号化とは，$\{0,1\}^*$ を任意の長さの 2 元系列の全体として，各 n に対して，写像 $\psi : \mathcal{X}^n \to \{0,1\}^*$ を定めることである（図 2）．x^n の符号長とは，その像の大きさ $|\psi(x^n)|$ のことである．データから符号長を対応させる関数を符号長関数と呼び，$L_\psi : \mathcal{X}^n \to \mathbf{R}^+$ と表す．

情報理論では，符号化に一種の条件を仮定する．それは**語頭条件**というものである．これは符号語が他の符号語の語頭と一致しないという条件である．これによって，符号語 $\psi(x^n)$ から元の x^n を誤りなく一意的に復号できることが保証される．

情報理論の分野では，ψ が語頭条件を満たすことの必要十分条件は，L_ψ によって以下のように完全に規定されることが知られている[1)]．

$$\sum_{x^n} 2^{-L_\psi(x^n)} \leq 1.$$

これは **Kraft の不等式**と呼ばれるものである．

そこで，語頭符号化関数 L_ψ が与えられたときに，P を各 x^n に対して

$$P(x^n) = 2^{-L_\psi(x^n)}$$

として定めよう．すると，Kraft の不等式により，

$$\sum_{x^n} P(x^n) \leq 1$$

が成り立つ．そのような P を X^n 上の劣確率分布と呼ぶ．逆に，確率分布 P が与えられていれば，

$$L(x^n) = \lceil -\log P(x^n) \rceil$$

と置くことにより，$L(x^n)$ は Kraft の不等式を満たすので，符号長関数になる．またそのような符号長関数を実現する符号化が実際に構成できることが知られている．この辺の事情については Cover and Thomas による情報理論の教科書[1)]などを参考にされたい．以上のことから，語頭符号化関数を一つ定めることは確率分布を定めることに他ならず，その意味で符号化と確率は表裏一体の概念であることがわかる．

以上の観点から式 (1) を再び眺めると，左辺は真の分布 P^* に関する平均符号長であり，その下限が P^*

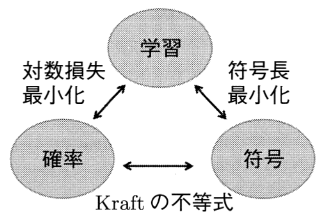


図3　学習・確率・符号化.

のエントロピーで与えられることがわかる．真の分布が事前に分かっていれば，$-\log P^*(x^n)$ の符号長で語頭符号化することにより，平均の意味で最も短い符号化が実現できるのである．

以上から，「学習」と「確率（分布）」が対数損失を通じてつながり，「確率」と「符号」が表裏一体の概念であることを見た（図 3）．さらに，「対数損失＝符号長」であるから，「学習を行うことは，符号長を短くする確率分布を見つけることに他ならない」という見方が成立するのである．

2. 正規化最尤符号化

さて，実際には P^* は分からないことが多いので，式 (1) の下限は容易には達成できない．まさにデータから P^* を学習することが必要になる．その際，全くの知識なしでは学習できない．そこで，データ生成分布の事前知識として確率モデルのクラスが与えられたとする．そのクラスを用いてできるだけ短い符号化を実現することを考えよう．

ここで，確率モデルのクラスとは，k を与えられた正の整数として，$\mathcal{C}_k = \{P(X|\theta) : \theta \in \Theta_k\}$ の形で書けるようなものとする．Θ_k は k 次元の実数値パラメータ空間であり，$P(X|\theta)$ は Θ_k の 1 点 θ により指定された確率モデルである．

$\mathcal{C}_k$ を用いた符号化法としては，2 段階符号化，ベイズ符号化，予測符号化など様々な方法が考えられる[3)]．ここでは，正規化最尤符号化[3)]を紹介しよう．

データ x^n からの θ の最尤推定量を $\hat{\theta}(x^n)$ で表す．すなわち，$\hat{\theta}(x^n) = \arg\max_{\theta \in \Theta_k} P(x^n|\theta)$ である．このとき，$P(X^n|\hat{\theta}(X^n))$ は $\mathcal{X}^n$ 上の確率分布をなさない．なぜなら，$\sum_{X^n \in \mathcal{X}^n} P(X^n|\hat{\theta}(X^n)) \geq 1$ となるからである．そこで，正規化して得られる

$$P_{NML}(X^n) = \frac{P(X^n|\hat{\theta}(X^n))}{\sum_{X^n \in \mathcal{X}^n} P(X^n|\hat{\theta}(X^n))}$$

を**正規化最尤分布**と呼び，各 x^n に対してこの確率分

布に基づく符号化の符号長：

$$
\begin{aligned}
&-\log P_{NML}(x^n)\\
&= -\log P(x^n|\hat{\theta}(x^n)) + \log \sum_{X^n \in \mathcal{X}^n} P(X^n|\hat{\theta}(X^n))
\end{aligned} \tag{2}
$$

を x^n の**正規化最尤符号長**と呼ぶ．

正規化最尤符号長の計算で難しいのは正規化項

$$
C_n(\mathcal{C}_k) = \sum_{X^n \in \mathcal{X}^n} P(X^n|\hat{\theta}(X^n))
$$

の計算である．これについては，$\mathcal{C}_k$ について，任意の x^n に対して最尤推定量 $\hat{\theta}(x^n)$ が真の θ の周りで中心極限定理を満たす場合については

$$
\begin{aligned}
&\log C_n(\mathcal{C}_k)\\
&= \frac{k}{2}\log\frac{n}{2\pi} + \log\int\sqrt{|I(\theta)|}d\theta + o(1)
\end{aligned} \tag{3}
$$

となることが知られている[4]．ここに，$I(\theta)$ は

$$
I(\theta) = \lim_{n\to\infty}\frac{1}{n}E_P^n\left[-\frac{\partial^2 \log P(X^n|\theta)}{\partial\theta_i\partial\theta_j}\right]
$$

で定義される **Fisher 情報量**と呼ばれる量であり，$|I(\theta)|$ はその行列式である．$o(1)$ は $n\to\infty$ につれて x^n に関して一様にゼロに収束する量である．

3. 確率的コンプレキシティ

正規化最尤符号長は符号長としてどれだけ短いものであろうか？ これをミニマックスリスクという規準を用いることによって評価しよう．

確率モデルのクラス $\mathcal{C}_k = \{P(X|\theta) : \theta \in \Theta_k\}$ が与えられたとし（簡単のため，Θ_k はコンパクトであると仮定する），$L(X^n)$ を X^n の語頭符号長関数であるとするとき，**ミニマックスリスク**[2,3]は以下で定められる．

$$
\min_L \max_{X^n}\left\{L(X^n) - \min_{\theta\in\Theta_k}(-\log P(X^n|\theta))\right\}.
$$

ここで，min は任意の語頭符号長関数に関してとられ，max はデータ列全体に関してとられるものとする．ミニマックスリスクは，モデルクラスによって定まる最小の対数損失 $\min_{\theta\in\Theta_k}(-\log P(X^n|\theta))$ に比べて語頭符号長の値をどこまで短くできるかをデータの最悪な場合について求めた値である．

ここでミニマックスリスクの min を達成する語頭符号長関数は式 (2) によって達成されることが容易に示される[2]．つまり，正規化最尤符号長はミニマックスリスクを達成する符号長としての最適性を満たすのである．この意味で，正規化最尤符号長 (2) はデータ列 x^n のモデルクラス $\mathcal{C}_k$ に相対的な最短符号長であると見なすことができる．

そこで，正規化最尤符号長 (2) を x^n の $\mathcal{C}_k$ に対する**確率的コンプレキシティ**と定義する[2~4]．これを $SC(x^n:\mathcal{C}_k)$ と表すと，正規化項について近似式 (3) が成り立つような場合については特に以下のように表される．

$$
\begin{aligned}
&SC(x^n:\mathcal{C}_k)\\
&= -\log P(x^n|\hat{\theta}(x^n))\\
&\quad + \frac{k}{2}\log\frac{n}{2\pi} + \log\int\sqrt{|I(\theta)|}\,d\theta + o(1).
\end{aligned} \tag{4}
$$

データの生成分布 P^* が既知のときに，Shannon 情報量 $-\log P^*(x^n)$ は x^n の最小記述長の意味をもった基本的情報量であった．一方，データの生成分布が未知のときには，確率的コンプレキシティを x^n の $\mathcal{C}_k$ を用いた最小記述長としての意味をもった基本的情報量であると見なすことができる．以下に見るように，学習問題は確率的コンプレキシティの k に関する最小化に帰着されるのである．

4. MDL 原理

モデルクラスの複雑さを表す指標である k はパラメータの次元を表している．データから最適なモデルの複雑さを求めることは学習の最も重要かつ代表的なテーマである．そのとき，「学習＝符号長最小化」という視点に立つと，確率的コンプレキシティ (4) を最小化するような k を求めることにより最適なパラメータの次元を決定できる．

このように，データが与えられたとき，これを表現するのに最もふさわしいモデルクラスの複雑さを決定する問題を**モデル選択問題**と呼ぶ．モデル選択規準として確率的コンプレキシティ最小（つまり符号長最小）を規準とする考え方を **MDL（Minimum Description Length）原理**（和訳：**記述長最小原理**）と呼ぶ[2,3]．

例として，ヒストグラムの学習をとりあげよう．今，与えられた k に対して，定義域を k 個の排反なセルに分割し，$\mathcal{X} = \cup_{i=1}^k C_i$，$C_i \neq C_j$ $(i \neq j)$ とする．各 $i \in \mathcal{X}$ について，$P(X \in C_i) = \theta_i$ をパラメータとし，$\sum_{i=1}^k \theta_i = 1$，$\theta_i \geq 0$ $(i = 1, \cdots, k)$ とする．そこで，

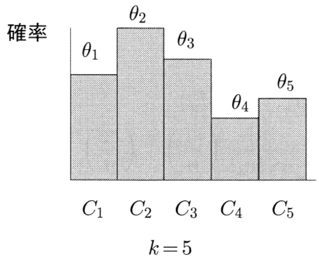


図 4　ヒストグラム.

$$P(X|\theta)=\sum_{i=1}^{k}\theta_i \mathbf{1}_{X\in C_i}/|C_i|$$

とする確率モデルを**ヒストグラム**と呼ぶ(図 4). $\mathbf{1}_{X\in C_i}$ は $X\in C_i$ のときに限って 1 の値をとり，その他の場合は 0 をとるような関数である. $|C_i|$ は C_i に属する X の総数を表す. つまり，ヒストグラムは各セルに出現する確率が θ_i で与えられるような構造をもった確率モデルである. 特に，各 i について，$|C_i|=1$ のとき多項モデルと呼ぶ. $k-1$ 次元実数値ベクトル $\theta=(\theta_1,\cdots,\theta_k)$ で指定されたヒストグラムのクラスを $\mathcal{C}_k$ と表す.

長さ n のデータ列 x^n が与えられたとして，$x\in C_i$ であるデータ数を n_i $(i=1,\cdots,k)$ とすると，θ の最尤推定量は $\hat{\theta}=(n_1/n,\cdots,n_k/n)$ であるから，確率的コンプレキシティは以下で与えられる.

$$\begin{aligned}&SC(x^n:\mathcal{C}_k)\\&=-\log\prod_{i=1}^{k}\left(\frac{n_i}{n|C_i|}\right)^{n_i}+\log C_n(k)\end{aligned}\tag{5}$$

ここに，$C_n(k)$ は次式で計算される量である.

$$C_n(k)=\sum_{n_1+\cdots+n_k=n}\frac{n!}{n_1!\cdots n_k!}\prod_{i=1}^{k}\left(\frac{n_i}{n}\right)^{n_i}.\tag{6}$$

ここで，近似式 (3) を用いると，

$$\log C_n(k)=\frac{k-1}{2}\log\frac{n}{2\pi}+\log\frac{\pi^{k/2}}{\Gamma(k/2)}+o(1)\tag{7}$$

で与えられる. $\Gamma(\cdot)$ はガンマ関数を表す. よって，MDL 原理によって最適なヒストグラムの複雑さ k を決定するには式 (5) に式 (7) の値を代入して得られる式を k の関数と見なして $SC(x^n:\mathcal{C}_k)$ を最小にする k を求めればよい.

5. クラスタリングへの応用

機械学習の最も実用的な問題であるクラスタリングの問題をとりあげ，ここに MDL 原理を適用してみよう.

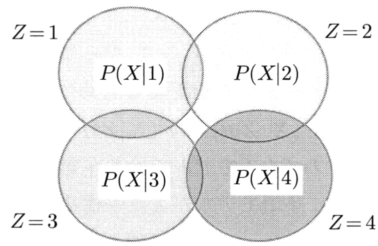


図 5　クラスタリングの確率構造.

クラスタリングとは，データを似たものどうしをまとめて幾つかのクラスターに分割することである. 今，データの発生が k 個のクラスターに分かれているとする. これを**有限混合モデル**を用いて表すとしよう (図 5). つまり，$P(X|Z)$ をクラスター Z の下でのデータ X の発生確率分布とし，$P(Z)$ をクラスター Z の生起確率分布として，各データ X が独立に確率モデル

$$P(X)=\sum_{Z=1}^{k}P(X|Z)P(Z)$$

に従って生成されると仮定する. Z は各データ X が生成されるクラスタを指定する変数であるが，X は Z に関して分布 $P(X|Z)$ の線形和で与えられる確率モデルから生成されると考えるのである. このような Z は実際には観測されないので，**潜在変数**と呼ばれる.

ここで，クラスターの数 k は未知である. クラスタリングの問題とは，1) 最適な k をデータ $x^n=x_1\cdots x_n$ から学習し，2) その k の下での各クラスターの分布 $P(X|Z)$ を学習し，3) 各データがどのクラスターから発生しているのか ($z^n=z_1\cdots z_n$) を推測することに帰着される.

今，$P(X|Z)$ の構造として**ナイーブベイズモデル**という構造を仮定しよう. データの次元は d であるとし，第 i 次元については $\{1,\cdots,M_i\}$ $(i=1,\cdots,d)$ の範囲に値をとり，各次元に関してデータの発生は独立であると仮定する. このとき，$P(X|Z)$ の構造は以下で与えられる.

$$\begin{aligned}&P(X^{(1)}=x^{(1)},\cdots,X^{(d)}=x^{(d)}|Z)\\&=\prod_{i=1}^{d}P(X^{(i)}=x^{(i)}|Z).\end{aligned}$$

ここで，$\pi_j=P(Z=j)$ $j\in\{1,\cdots,k\}$ と $\sigma_{ijl}=P(X^{(i)}=l|Z=j)$ $(i\in\{1,\cdots,d\},\ j\in\{1,\cdots,k\},$

$l \in \{1, \cdots, M_i\}$) をパラメータとする．ただし，$\sum_{j=1}^{k} \pi_j = 1$, $\sum_{l=1}^{M_i} \sigma_{ijl} = 1$ なる制約がある．このようなパラメータを θ としてまとめて表し，X と Z の同時分布を $P(X, Z|\theta) = P(X|Z)P(Z)$ と表す．つまり，ナイーブベイズモデルでは潜在変数 Z の方向と各次元の方向に多項モデルが入れ子構造をなしている．

データ列 $x^n = x_1 \cdots x_n$ に対して，何らかの方法で各 x_i に対応するクラスターを z_i と決めたとして，クラスター列を $z^n = z_1 \cdots z_n$ と記すとする．このとき，(X^n, Z^n) の正規化最尤分布は

$$P_{NML}(X^n, Z^n|\mathcal{M}_k) = \frac{P(X^n, Z^n|\hat{\theta}(X^n, Z^n))}{C_n(k, M_1, \cdots, M_d)}$$

のように計算できる．ここに，

$$C_n(K, M_1, \cdots, M_d) = \sum_{X^n, Z^n} P(X^n, Z^n|\hat{\theta}(X^n, Z^n)).$$

今，データ系列とこれに対応するクラスター系列の実現値 (x^n, z^n) が与えられたとする．ここで，z^n の良し悪しを確率的コンプレキシティを用いて評価したい．h_j を z^n の中でクラスター j に属するものの数とし，f_{ijl} を x^n の中で，i 番目の要素が値 l をとり，対応するクラスターが j であるものの数であるとすると，正規化最尤分布の分子は次式で計算できる．

$$P(x^n, z^n|\hat{\theta}(x^n, z^n)) = \prod_{j=1}^{k} \left(\frac{h_j}{n}\right)^{h_j} \prod_{i=1}^{d}\prod_{l=1}^{M_i} \left(\frac{f_{ijl}}{h_j}\right)^{f_{ijl}}.$$

また，分母は次式で与えられる．

$$C_n(k, M_1, \cdots, M_n) = \sum_{h_1+\cdots+h_k=n} \frac{n!}{h_1! \cdots h_k!} \prod_{j=1}^{K} \left(\frac{h_j}{n}\right)^{h_j} \times \prod_{i=1}^{d} C(M_i, h_j). \quad (8)$$

ここに，

$$C(M_i, h_j) = \sum_{f_{ij1}+\cdots+f_{ijM_i}=h_j} \frac{h_j!}{f_{ij1}! \cdots f_{ijM_i}!} \prod_{l=1}^{M_i} \left(\frac{f_{ijl}}{h_j}\right)^{f_{ijl}}.$$

この場合は，$C_j(1, M_1, \cdots, M_d) = \prod_{i=1}^{d} C(M_i, j)$ ($j = 1, \cdots, n$) として Kontkanen *et. al.* により，各 M と j に対して $C(M, j)$ が以下の漸化式が満たすこと，

$$C(M+2, j) = C(M+1, j) + (j/M)C(M, j)$$

および $C_n(k, M_1, \cdots, M_d)$ が以下の漸化式を満たすことが示されている[5]．

$$C_n(k, M_1, \cdots, M_d) = \sum_{r_1+r_2=n} \frac{n!}{r_1! r_2!} \left(\frac{r_1}{n}\right)^{r_1} \left(\frac{r_2}{n}\right)^{r_2} \times C_{r_1}(1, M_1, \cdots, M_d) \times C_{r_2}(k-1, M_1, \cdots, M_d).$$

これらの漸化式を用いると式 (8) は計算量 $O(n^2 k)$ で計算できる．

このようにして正規化最尤分布の符号長が求められると，最良のクラスターの数 $\hat{k}$ とクラスタリングの結果 $\hat{z}^n$ は確率的コンプレキシティを最小にするものとして以下のように求めることができる．

$$(\hat{k}, \hat{z}^n) = \arg\min_{(k, z^n)} \{-\log P(x^n, z^n|\hat{\theta}(x^n, z^n)) + \log C_n(k, M_1, \cdots, M_d)\}.$$

以上が，MDL 原理に基づく最適なクラスタリングの決定方法である．

上では，幾つかのクラスター系列 z^n の候補が与えられた下で最適な $\hat{k}$ と $\hat{z}^z$ を選ぶ方法を与えた．z^n の候補を作り出す方法としては様々な方法が考えられる[5]．例えば，**EM アルゴリズム**という方法を用いることができる．これは各クラスターにアサインされる事後確率を計算し，事後確率が最大になるクラスターを割り当てる方法である．

本節で示したクラスタリング手法をグラフのクラスタリングに適用した例を示す（図 6）．グラフクラスタリングとは，ノードの関係性をグラフで示したときに，より関係して結びついているノードの群を一つのクラスターとして見なす，ノードのクラスタリングのことである．図では 4 つのクラスターに分かれていることが示されている．このようなクラスタリングはソーシャルネットワークやブログなどのリンク関係をもとに幾つかのコミュニティに分割することに相当する．これ

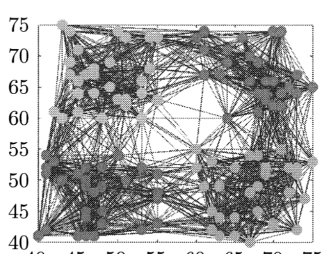


図 6　グラフクラスタリング．

によって，幾つのコミュニティが存在するか？ といった構造を知ることができる．

6. データマイニング：異常検知への適用

前節のクラスタリングでは似たようなものどうしをクラスターとしてまとめあげることを問題にしていた．これは，見方を変えると，どこのクラスターにも属さない外れ値データを検知することにも通じる．これを**外れ値検出**と呼んでいる．

大量のデータからの外れ値は，例えば，データがネットワークへのアクセスログである場合，不正侵入や情報漏えいなどの犯罪行為に関連したデータである可能性が高い．また，コンピュータ機器の性能データ（例えば，CPU ロードなど）である場合には，機器の障害に関係するデータである可能性が高い．そのように，外れ値検出は，セキュリティ分野における犯罪の検知や機械系の障害検知などに現実に応用できるのである[7]．

また，実際に外れ値検出を行うには，基になるクラスタリングを，データが与えられる毎にオンラインで行わなければならない．さらに，データの生成源の性質が時間とともに変化しても適応的にクラスタリングしていかねばならない．特に，後者については，最適なクラスターの数が時間とともに変化するような場合をも扱う．そのように時間的に変化するモデルの系列をトラッキングする問題は，最近，**動的モデル選択**と呼ばれる文脈の中で研究されている[6]．これはデータから確率モデルの構造的変化を扱う大きなトピックである．

外れ値検出や，確率モデルの構造的変化の検出などを通じて，大量データの中に含まれる異常を検出する問題を総称して**異常検知**と呼んでいる[7]．異常検知問題は大量データからの知識発見技術（これを**データマイニング**と呼ぶ）の中でも現実的な要請の高いテーマの一つであり，セキュリティやリスク管理などビジネス的にも極めて広い範囲に応用可能性をもっている．異常検知問題の詳細について興味のある方は拙著[7]をも参照されたい．

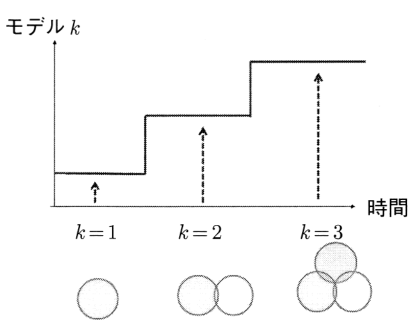


図7　動的モデル選択．

7. おわりに

学習の基本問題としてモデル選択問題から始まり，クラスタリング問題，そして異常検知問題といったデータマイニングのビジネス問題まで話が及んだ．これらの技術は今や，製造業・流通業・サービス業などの産業界で大きなニーズをもって迎えられている．ここで，その根っこにあるのは「確率」という概念であることを忘れてはなるまい．事実，学習やデータマイニングの分野では，「確率」はデータを語るための共通言語である．本稿では，「確率」と「符号」が表裏一体の概念として結びつき，それゆえに生まれる「符号長最小化」の考え方が「学習」の本質に結びついていることを示した．現実の学習問題がどんなに複雑に，かつ大規模になろうとも，上記の考え方はそれを解くための一つの指針となり得るであろう．MDL 原理の考え方をより深く理解したい読者は創始者である Rissanen がまとめた最近の著書[3]や Grünwald らが最近の話題をまとめた著書[2]を参考にされたい．

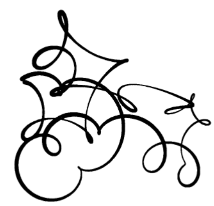
参考文献

1) T.M. Cover and J.A. Thomas. *Elements of Information Theory*, Wiley-Interscience, 1991.
2) P. Grünwald, J. Myung, M.A. Pitt. *Advances In Minimum Description Length: Theory And Applications*. The MIT Press, 2005.
3) J. Rissanen. *Information and Complexity in Statistical Modeling*, Springer, 2007.
4) J. Rissanen. Fisher information and stochastic complexity. *IEEE Transaction on Information Theory*, vol.**42**, 1, pp:40–47, 1996.
5) P. Kontkanen, P. Myllymäki, W. Buntine, J. Rissanen, and H. Tirri: An MDL framework for data clustering. In P. Grünwald, I. Myung, and M. Pitt, editors, Advances in Minimum Description Length: Theory and Applications. The MIT Press, 2005.
6) K. Yamanishi and Y. Maruyama: Dynamic model selection with its applications to novelty detection. *IEEE Transactions on Information Theory*, vol.**53**, 6, pp:2180–2189, June, 2007.
7) 山西健司：『データマイニングによる異常検知』，共立出版，2009.

（やまにし・けんじ，東京大学大学院情報理工学系研究科）

数理ファイナンスに現れる確率過程

確率過程を用いた資産価格過程のモデル化



関根 順

1. 序

数理ファイナンス分野では，大雑把に言って，「確定的でない (non-deterministic)，不確実な (uncertain) 金融市場でどのような意思決定を行うか？」に関する考察が主要になる．

例を挙げてみよう．まず，時刻 t での n 種類の危険資産価格[*1]を

$$S_t := (S_t^1, \ldots, S_t^n)$$

で表す．また，時刻 0 を開始時刻（あるいは現在）と見なし，期末時刻 $T(>0)$ を一つ固定して危険資産の価格過程を

$$S := (S_t)_{t \in [0,T]}$$

と表すことにする．以下は，ファイナンス分野での重要かつ基本的な問題である．

問題 1 （金融派生商品の価格付け）満期日 T での価値が所与のペイオフ関数 f を用いて

$$F := f(S) \tag{1}$$

で与えられるような金融商品[*2]を考える．この金融商品の任意の時刻 $t \in [0,T]$ での“適正な”理論価格を考察せよ．

問題 2 （最適投資）時刻 0 に所与の元手 x 円で n 種類の危険資産に対して投資を開始する．投資家は，時刻 0 から T まで危険資産の保有量を調整しながら刻々と運用を続ける．時刻 T での富を“最適化”するような運用の仕方を考察せよ．

これらの問題に対して，危険資産の価格が確定的でないとし，資産価格過程を確率過程として記述するのが通常のアプローチである．すなわち，現時点で将来の資産価格をピンポイントで知ることはできない，ただしその確率分布はわかっている，と見なすわけである．

本稿では，数理ファイナンス分野で，危険資産価格過程 S の数理モデルとして，どのような確率過程がどんな目的意識で用いられているのか，調べられているのか，その一端を紹介したい．

理論的見地から問題を論じようとする際は，例えば，S を半マルチンゲールの範疇から採り[*3]，なるべく一般的な立場から考察する．一方，実務への応用を意識したり，具体例を用いて経済学的な解釈を結果に加えたりする際は，数値的・明示的な解析が可能なより限定的で特別なクラスに S を

*1) ここで，危険資産とは，例えば，様々な銘柄の株式，外国通貨，物品（コモディティ：石油など天然資源，金など天然鉱物，麦・とうもろこしなど穀物），（金利変動や倒産のリスクを考慮した）債券，などを総称した言い方である．

*2) 実例としては，$n=1$ として $F=\max(S_T-K,0)$（ヨーロピアンコールオプション，K は行使価格と呼ばれる正定数），$F=\max\left(\frac{1}{T}\int_0^T S_t dt - K, 0\right)$ （エイジアンコールオプション），$F=S_T-\inf_{t\in[0,T]} S_t$ （ルックバックコールオプション），などが挙げられる．

*3) この範疇では確率解析（stochastic calculus）の助けを借りた解析が可能である．

限ることになる．いずれにせよ，目的に応じて現実の金融市場を理想化・単純化して作った山ほどのモデルがあるというのが現状である．自然現象とは異なる，決して綺麗ではない金融市場が対象なので，すべてを説明できる美しい統一的なモデルのようなものは存在しないように思われる．本稿では，山ほどあるモデルを羅列するのではなく，むしろ，確率過程モデルを作る際の研究者・実務家の問題意識をいくつか紹介することを狙いたい．

また，上述の2つの問題を見ても想像できるように，決して確率過程モデル S の構成や解析だけが数理ファイナンスの問題ではないことも強調しておきたい．本稿では触れないが，例えば，問題1で何をもって「適正」と呼ぶのか，あるいは，問題2でどんな意味・基準で「最適化」を行うのか，また，資産価格モデルの不確実さ (uncertainty)，例えば価格の真の確率分布がわからない状況を，どう取り扱うのか？ これらすべてが大変重要な問題点である．そして解答を与えるには，数学のみならず，関連した分野：経済学・工学・統計学などの智慧も必要になる部分である．

さて，本稿の組み立ては以下の通りである．まず次節で，最も基本的ないわゆるブラック・ショールズ・マートンモデルを導入する．続いて3–4節で，上記の問題1で重要になる，リスク中立確率 $\mathbb{Q}$ の下でのモデル化を解説する．この場合は，資産価格のランダムネスの強度を表すボラティリティをどうモデル化してゆくかが重要になる．最後の5節では，少ない分量ではあるが，上記の問題2で重要になるような，実世界確率 $\mathbb{P}$ の下でのモデル化に若干言及する．この場合は，ボラティリティのみならず資産価格過程の期待成長率項のモデル化も行わないといけない．

■ 2. 指数ランダムウォークモデル・指数ブラウン運動モデル ■

まずは，ファイナンス分野での基本中の基本とも言えるモデル，いわゆるブラック・ショールズ・マートンモデルを紹介しよう．その前段階として，より簡単な，離散時間

$$\mathcal{T} := \{0, \delta, 2\delta, \ldots, N\delta\}$$

（δ は時間の刻み幅を表す定数）上の 1 次元危険資産価格過程

$$S := (S_t)_{t\in\mathcal{T}}$$

のモデルの紹介から始めよう．このモデルでは，価格の変動や売買取引はすべて離散時刻 $\mathcal{T}$ でのみ起こると考えているわけである．

さて，時刻 t から $t+\delta$ の間の危険資産の収益率（リターン）を

$$R_t := \frac{S_{t+\delta} - S_t}{S_t}$$

と表す．これはすなわち，時刻 t に危険資産を 1 単位保有したとき，$t+\delta$ でどれだけの収益が得られるか割合で表したものである．以下を仮定する．

(1) 各期間の収益率は独立で同分布な確率変数である．

(2) 平均と分散がそれぞれ

$$\mathbb{E}[R_t] = \mu\delta, \quad \mathbb{V}[R_t] := \mathbb{E}(R_t - \mathbb{E}R_t)^2 = \sigma^2\delta$$

と表されている．ここで，μ は年率あたりの期待収益率，σ^2 は年率で表した収益率の分散を表している．また，σ（= 年率で表した収益率の標準偏差）は**ボラティリティ**とも呼ばれる．

この場合，

$$S_t = S_0 \prod_{i=1}^{t/\delta}(1+R_i), \quad t \in \mathcal{T}$$

と独立同分布の確率変数の積で書かれることになる．あるいは

$$S_t = S_0 \mathrm{e}^{Z_t}, \quad t \in \mathcal{T},$$

ただし

$$Z_t := \sum_{i=1}^{t/\delta} z_i, \quad z_i := \log(1+R_i)$$

とも表される．例えば，上記の (1)–(2) を満たす簡単な例として，

$$R_t = \begin{cases} \mu\delta + \sigma\sqrt{\delta} & \text{確率 } 1/2 \\ \mu\delta - \sigma\sqrt{\delta} & \text{確率 } 1/2 \end{cases}$$

と採ってみよう．この場合，独立な確率変数 z_i の和で表される $(Z_t)_{t\in\mathcal{T}}$ は，各時刻で 1/2 ずつの確率で増減を繰り返す，ランダムウォークと呼ばれる確率過程になっている．すなわち，危険資産価格はランダムウォークの指数関数で表されているわけである．

次に，期末時刻 $T = N\delta$ を固定して，時間の刻み幅 δ を 0 に近づけてゆくことを考える．売買取引が頻繁に行われており，その時間間隔が微小な状況を考えていると言ってもいいかもしれない．$\delta \to 0$ の極限として得られるのが

$$S_t = S_0 e^{\sigma w_t + \left(\mu - \frac{1}{2}\sigma^2\right)t}, \quad t \in [0, T] \qquad (2)$$

と表される指数ブラウン運動である．ただしここで，$w := (w_t)_{t\in[0,T]}$ はブラウン運動と呼ばれる確率過程であり，S に含まれるノイズ項（ランダムネスの源）を表現している．そしてこれが，ファイナンス分野で**ブラック・ショールズ・マートンモデル**（以下 BSM モデルと略記）[*4)] と呼ばれている連続時間危険資産価格過程モデルである．

図 1 にブラウン運動をコンピューター上でシミュレートして描いた BSM モデル (2) のサンプルを載せた．

(2) 式に伊藤の公式を適用して，微分形で表すと

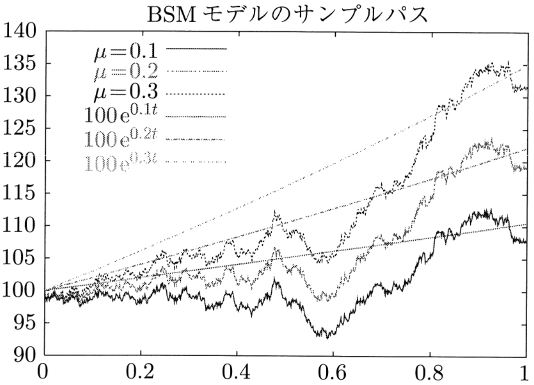


図 1　シミュレーションで描いた BSM モデル（期待収益率 μ を 3 通りの値に変化させてシミュレーションを行った．一緒に $\sigma = 0$ と置いた $S_0 e^{\mu t}$ のグラフも描いている）．

*4)　ブラック・ショールズ・マートンが，有名な金融派生商品の価格付けに関する研究を行った際に採用した危険資産価格過程モデルであることから，この呼び名が付いている．

$$\frac{dS_t}{S_t} = \mu dt + \sigma dw_t \qquad (3)$$

とも表される．この式がより解釈し易いかもしれない．微小な時間 dt が経過するとき，(3) 式は

$$\frac{S_{t+dt} - S_t}{S_t} \approx \mu dt + \sigma(w_{t+dt} - w_t) \qquad (3')$$

のように S_t が時間発展してゆくと解釈される．またここで，ブラウン運動の増分

$$w_{t_i} - w_{t_{i-1}}, \quad (t_i := i dt,\ i = 1, 2, \ldots,)$$

が互いに独立で，$w_{t_i} - w_{t_{i-1}}$ は平均 0 で分散が $t_i - t_{i-1}$ の正規分布していることに注意する．したがって，(3)–(3') 式は，BSM モデルの以下の性質を意味している．

BSM-1　収益率（リターン）が各々の期間で独立である（累積リターン $\int_0^t dS_u/S_u$ が独立な増分を持っている）．

BSM-2　収益率は平均 μdt，分散 $\sigma^2 dt$ の正規分布している．

本節の冒頭に述べたように，BSM モデルはファイナンス分野での基本中の基本のモデルである．今日では，すべての研究者（数学者，経済学者，工学者 etc.），金融実務家（クオンツと呼ばれる商品開発・解析のスペシャリスト，トレーダー etc.）が共有している「言語」といっても過言でないくらいである．もちろん，誰も現実の資産価格が上記の BSM-1, BSM-2 を満たしているなどとは信じていない．極めて単純化・理想化されたモデルの一つであるわけだが，ベースラインとして，解釈やコミュニケーションの道具として，研究や金融商品の開発現場で用いられ続けている．

■ 3.「リスク中立確率 $\mathbb{Q}$ の下でのモデル化」＝「ボラティリティのモデル化」 ■

前節の脚注 4 の中で触れた，ブラック・ショールズ・マートンによる金融派生商品価格付けの結果の骨子を紹介しよう．そこでは，**リスク中立確率** $\mathbb{Q}$ と呼ばれる，現実世界のモデルを記述するのに想定されている実世界確率 $\mathbb{P}$ とは異なる確率を

導入するところがポイントになっている．リスク中立確率の存在や構成法については本稿では触れないが，例えばBSMモデル (2)–(3) をこの確率の下で記述すると

$$S_t = S_0 \mathrm{e}^{\sigma w_t^{\mathbb{Q}} + \left(r - \frac{1}{2}\sigma^2\right)t},$$

あるいは

$$\frac{dS_t}{S_t} = r dt + \sigma dw_t^{\mathbb{Q}} \tag{4}$$

と表現される．ここで，定数 r は安全運用の利率[*5)]を表しており，

$$w_t^{\mathbb{Q}} := w_t + \frac{\mu - r}{\sigma} t$$

はリスク中立確率 $\mathbb{Q}$ の下でのブラウン運動である．(3) 式と (4) 式を比べてみると，実世界確率 $\mathbb{P}$ の下では危険資産価格の期待収益率 μ に当っていた部分がリスク中立確率 $\mathbb{Q}$ の下では安全運用の利率 r に等しくなっており，これを「リスク中立化」と呼んでいるわけである．

さて，以下に結果の骨子を述べよう．

金融派生商品の適正価格　資産価格過程[*6)] S を用いて，満期時 T での価値が (1) で表される金融商品を考える．この金融商品の時刻 $t \in [0, T]$ での適正価格[*7)] は

$$V_t := \mathbb{E}_t^{\mathbb{Q}}[\mathrm{e}^{-r(T-t)}F]$$

である．ただしここで，$\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ はリスク中立確率 $\mathbb{Q}$ に関する時刻 t での条件付期待値を表す．

例えば，脚注2に挙げたヨーロピアンコールオプション $F = \max(S_T - K, 0)$ の場合は

$$V_t := C(t, S_t, T, K, r, \sigma),$$

ただし

$$C(t, x, T, K, r, \sigma) = x\Phi(d_1) - K\mathrm{e}^{-r\tau}\Phi(d_2),$$

$$\Phi(d) := \int_{-\infty}^{d} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-y^2/2} dy,$$

$$d_1 := \frac{1}{\sigma\sqrt{\tau}} \left\{ \log\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau \right\},$$

$$d_2 := d_1 - \sigma\sqrt{\tau}, \quad \tau := T - t,$$

とより明示的に計算される．これは，ブラック・ショールズのコールオプション価格公式と呼ばれている．

上記の結果の重要な帰結の一つは，ボラティリティ σ こそが，金融派生商品の価格付けの際，推定・決定が必要となる唯一のパラメータである，という点である[*8)]．これは，理論的な適正価格がリスク中立確率に関する期待値で表現できることに拠っている．そのため，派生商品価格付けの目的においては，資産価格過程のボラティリティをどのように推定・決定してゆくかが重要課題となる．さらに，やはり定数ボラティリティ σ を持つBSMモデルでは不十分であり，発展モデルの研究が現在でも活発に続けられている現状である．

例えば，以下の市場データを用いた観察を見てみよう．まず，ヨーロピアンコールオプションのインプライドボラティリティを導入しておく．これは，売買されているヨーロピアンコールオプションの市場価格からブラック・ショールズ公式を用いて逆算されたボラティリティの値のことである．すなわち，満期日が T で行使価格が K のヨーロピアンコールオプションが売買されており，現在時刻 t での市場価格が $C_{\mathrm{mkt}}(t, T, K)$ で付いていたとしたとき，

$$C_{\mathrm{mkt}}(t, T, K) \quad (\text{: 市場価格})$$
$$= C\left(t, S_t, T, K, r, IV_t(T, K)\right) \quad (\text{: 理論価格})$$

を満たす値 $IV_t(T, K)$ が，時刻 t での市場価格から算出された**インプライドボラティリティ**である．表1に，日経225（主要株式銘柄225個の株価の重み付き平均）に関するヨーロピアンコールオプションのインプライドボラティリティの値を，2007年4月4日の市場価格から算出したものを

*5)　簡単には，銀行に預金したときに付く利率と考えればよい．

*6)　BSMモデルが例になっているような，ある一定の仮定を満たすものとする．

*7)　ここで「適正」とは「無裁定」という意味である．これ以上の説明は割愛する．

*8)　コールオプションの価格公式を見ると明らかであろう．現在時刻 t，現時点の資産価格 S_t，満期日 T，定数 K（行使価格と呼ばれる），預金金利 r，は推定の必要がない所与の値である．

表1　インプライドボラティリティマトリックス．

	9	37	65	100	163	254	345
215			17.1				
210			16.8				
205		17.8	16.3				
200		16.4	16.2	16.6	16.8		
195	25.6	16.1	16.3	16.8	16.6	15.6	
190	19.9	16.1	16.0	16.1	16.6	16.1	
185	18.0	15.0	16.2	16.3	16.6	16.1	
180	16.5	15.4	16.8	16.4	16.6	15.3	
175	15.5	16.2	17.4	16.6	15.5	15.2	
170	18.9	16.6		15.1		14.6	
165	31.4	19.0		17.6			
160	25.7						
155	9.7						
150	12.5						
145			6.7				17.1

記してみた*9)．

表の見方は以下の通りである．

- 一番左の行の数値が行使価格を表す（単位100円）．
- 一番上の列の数値が満期までの期間（9, 37, 65, 100, 163, 254, 345; 単位：日）を表す．
- 各数値は(満期, 行使価格) = (T, K) のインプライドボラティリティ値 $IV_0(T, K)$（単位：パーセント，時間の単位は年にとって計算）を表す．

満期 T，行使価格 K 毎に異なるインプライドボラティリティ IV の値が算出されていることが見てとれる．表1の値をプロットし，適当に補間して描いたものが図2である（ボラティリティサーフェスと呼ばれる）．

もし，日経225平均株価がBSMモデルにしたがって変動しているならば，ボラティリティサーフェスはフラットな表面になるはずである．対偶をとって言うと，市場で観測されるようなフラットでないボラティリティサーフェスは，BSMモデルのような簡単なモデルでは再現できないということである．実務的には，

「市場価格」≈「理論価格」

*9)　ウェブサイト：Pan Rolling（http://www.panrolling.com/）で閲覧できるデータを用いた．

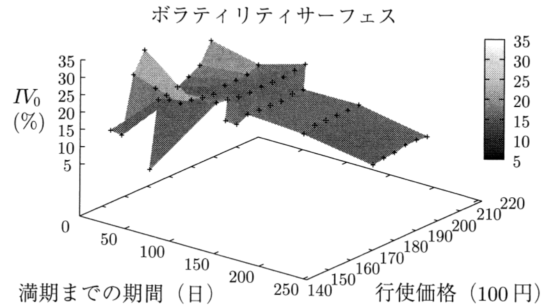


図2　ボラティリティサーフェス．

がある程度満たされるような危険資産価格過程モデルが要請されている*10)．

次節で，いくつか発展モデルを紹介しよう．

4. BSMモデルの発展

リスク中立確率 $\mathbb{Q}$ の下で記述されたBSMモデル(4)の発展形は，例えば，

$$\frac{dS_t}{S_t} = rdt + \sigma_t dw_t^{\mathbb{Q}} \tag{5}$$

（ここで，ボラティリティ過程 $(\sigma_t)_{t\geq 0}$ をどうモデル化するかがポイント），あるいはより一般的に

$$\frac{dS_t}{S_t} = rdt + dM_t$$

（ここで，$\mathbb{Q}$-マルチンゲール $(M_t)_{t\geq 0}$ をどうモデル化するかがポイント）と書かれる．いくつか例を紹介しよう．

4.1 局所ボラティリティモデル

(5)式で

$$\sigma_t = a(t, S_t)$$

と2変数関数 $a : [0, \infty) \times [0, \infty) \to [0, \infty)$ を用いて構成したモデルを総称して局所ボラティリティモデルと呼ぶ．現在時刻 $t = 0$ において，所与のボラティリティサーフェスに完全に合致させる（すなわち，時刻 $t = 0$ において，コールオプションの「市場価格」と「モデルを用いた理論価格」が常に合致する）には，このクラスに限って

*10)　一方で，闇雲に市場価格と理論価格を合わせるようにモデルパラメータを調整していくことの危険性も，特に今回の世界的金融危機以降，指摘されている．

考察すれば十分であることが知られている．

4.2 時系列構造の入った経路依存モデル

文献 3) では，ボラティリティ過程の過去への依存度を確率微分方程式を用いて簡潔にモデル化している．これは統計分野の時系列解析の手法にヒントを得たものである．単純な例は以下である．(5) 式で，ボラティリティ過程が

$$\sigma_t = a(Y_t) \tag{6}$$

と別の確率過程 Y を用いて記述されているとする．ただし，ここで

$$Y_t := \log S_t - \int_{-\infty}^{t} (\log S_u) k e^{-k(t-u)} du$$

と定義されている．上式の右辺は時刻 t での対数価格とそれ以前の対数価格 $(\log S_u)_{u\in(-\infty,t)}$ の重み付き平均の差である．パラメータ $k > 0$ の値が小さいほど，過去の価格に依存した形でボラティリティ σ_t が決まっていることになる．

一般に，(6) 式のように，ボラティリティ過程が別の確率過程を用いて表されているモデルを確率ボラティリティモデルと呼んでいる．どのような Y をまた関数 a を選ぶかはたくさんの選択があり，膨大な研究が今なお活発に行われている．

4.3 ボラティリティサーフェスモデル

最近の研究の流れで興味深く思われるものの一つとして，時間と共に確率的に変動してゆくボラティリティサーフェスを直接モデル化している研究を挙げておこう．

これは，本稿ではすべて割愛してしまった金利過程モデルの研究の中で現在主流を占めている，イールドカーブモデルの影響を大きく受けている．興味のある読者の方は，例えば文献 1) を参照していただきたい．関連研究のレビューや著者自身の最新モデルの紹介がある．

■ 5. 実世界確率 $\mathbb{P}$ の下でのモデル化 ■

最後に，実世界確率 $\mathbb{P}$ の下で行う危険資産価格過程モデル S の構成について若干言及したい．3–4 節で解説したような金融派生商品の価格付け (1 節の問題 1) においては，リスク中立確率 $\mathbb{Q}$ の下で記述されたモデルしか意識されないことも多いのだが，例えば 1 節の問題 2 を扱う際などは，やはり，実世界確率 $\mathbb{P}$ の下で記述されたモデルが必要になる．(3) 式で書かれる BSM モデルやその発展，

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dw_t$$

を考えることになるが，特に，期待収益率項 μ の推定が困難であることが各所で指摘されている．そこで例えば，工学の世界で盛んに用いられているフィルタリングの手法の応用が考えられるかもしれないが，この場合も μ の事前分布や構造の同定を行う必要があり，これらは決して易しい問題ではないように思われる．

μ_t の構造の同定をある程度理論的に行うのに，経済学的な均衡アプローチを採ることも考えられる．文献 2) は，初学者には難しいと思われるが，数学の立場から経済学の均衡などにも言及した格調高い数理ファイナンスの教科書として知られている．また，拙論文 4) にも簡単な均衡アプローチを用いた考察が載っている．

参考文献

1) Carmona, R. (2009). "Market models for European options: dynamic local volatility and tangent Lévy models", *Workshop on the Pricing and hedging of environmental and energy-related financial derivatives*, (4–9 Dec. 2009, Institute for Mathematical Sciences, National University of Singapore) における講演資料，http://www.ims.nus.edu.sg/Programs/financialm09/activitieswk2.php にてダウンロード可能．

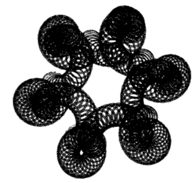
2) Föllmer, H. and A. Schied. "Stochastic Finance, An Introduction in Discrete Time". Walter de Gruyter, Berlin, New-York, 2002.

3) Hobson, D.G. and L.C.G. Rogers (1998). Complete models with stochastic volatility. *Mathematical Finance*, **8**(1) 27–48.

4) 関根順 (2008). 均衡価格過程から導出されるリスクプレミアムについて (「フィナンシャルテクノロジーの過去・現在・未来」, 創立 20 周年記念論文集，三菱 UFJ トラスト投資工学研究所，http://sites.google.com/site/junsekine/works からダウンロード可能).

(せきね・じゅん，京都大学経済研究所)

経済学と確率



中村　恒

要旨　本稿では最近の経済学における連続時間の確率解析の手法を使った分析を紹介する．

1. はじめに

経済学では，効用あるいは収益を最大化するために，数量に限りのある資源を消費・投資行動を通じてどのように配分し効率的に利用すべきか，という問題を分析する．特に現実経済では，経済主体が消費・投資の意思決定を行う時点では，その意思決定から将来どのような効用・収益が得られるかが不確実（uncertain）であることが一般的である．そこで，経済学では不確実性の下での最適な消費・投資を分析し，その結果実現される市場均衡での価格やポートフォリオの動向を観察する．不確実性の下で経済主体の取り得る最善の行動は，消費・投資行動の成果が将来どのように分布するのかを確率的に分析し，その分布自体を価値判断し選択することである．つまり意思決定の選択対象は確率過程である．不確実な将来の見通しは，成果の集合の上に定義される確率過程によってモデル化される．そしてその確率過程の価値を評価するために，経済学では期待効用アプローチと呼ばれる分析方法がしばしば使用されてきた．そこでは，確率過程の価値を，それのもたらす効用（あるいは収益）の期待値――これを**期待効用**あるいは誤解がない限り単純に**効用**と呼ぶ――によって評価する．特に期待効用は典型例として以下のように定式化される．以下では本節を通じて，時間は連続で終期 $T>0$ を所与として有限な期間 $[0,T]$ とし，$(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\le t\le T},P)$ は通常の条件を満たすフィルター付き確率空間で，情報系（filtration）は 2 次元のブラウン運動 $\{B_t\}_{0\le t\le T}$ によって生成されるとする．また，消費過程 $\boldsymbol{c}=\{c_t\}_{0\le t\le T}$ をある正則条件を満たす $\{\mathcal{F}_t\}_{0\le t\le T}$ に適合な非負実数値確率過程とする．このとき期待効用 U_t は，定数パラメータ β,δ を与件として，終期 T での効用を $U_T=0$ とした場合，

$$U_t=\begin{cases} E_t\left[\displaystyle\int_t^T e^{-\beta(s-t)}\frac{c_s^{1-\delta}}{1-\delta}ds\right] \\ (\delta\neq 1\text{ のとき}), \\ E_t\left[\displaystyle\int_t^T e^{-\beta(s-t)}\log(c_s)ds\right] \\ (\delta=1\text{ のとき}) \end{cases} \tag{1.1}$$

と表される．なお，β は時間選好を表し，現在の消費からの効用にどれだけのウェイトをおくのかを示し，δ は危険回避度を表し，どれだけ消費の不確実性を回避したいと思っているのかを示す．期待効用アプローチでは，効用は状態に関して分離可能である上に，時間に関しても分離可能であることが仮定される．また，情報構造は客観的（objective）に外から与えられると仮定される．このような期待効用アプローチは，連続時間モデルでは，文献 1) や文献 12) 等での条件付き請求権（contingent claim）モデルや文献 5, 6) での均衡資産価格モデルなどで使用されている．

しかし，このアプローチに関して現実経済を説明する上でいくつかの深刻な問題点が指摘されてきた．代表的な問題点として 2 つ挙げることができる．第一は情報構造に関する問題であり，第二は効用関数の性質に関する問題である．最近 10 年の間に，これらの問題を解決するために連続時間の確率解析の手法を用いた経済分析が進んでいる．

まず前者について簡潔に触れておく．上述のように期待効用アプローチでは情報構造が客観的に外から与えられることを前提とする．しかし，個々の経済主体は将来に対して異なった期待（あるいは確率分布）を

主観的（subjective）に形成すると考えるのが現実的な場合が多い（なお客観的確率，主観的確率の詳しい解説は文献 11) 等を参照）．また，情報は経済全体に分散し，ある経済主体は他の経済主体よりも限られた情報の下で意思決定することもしばしばで，しかもある経済主体が戦略的に情報構造を操作することもある．そうした状況を鑑みて，経済学では以前から，情報の非対称性やモラルハザード，モデルの不確実性（model uncertainty）等の情報問題が存在するケースでの意思決定の分析が発展してきた．ただ，これまでは素朴な有限離散時間モデル（典型的には 2,3 期間モデル）での分析が多く，現実への応用は限定的であった．しかし，最近 10 年の間に，それらの情報問題に関して連続時間の確率解析の手法を使って動学的に複雑な情報構造をモデル化し，特に金融経済学の分析に多く応用されている．例えば，文献 16) では無限期間の連続時間モデルで，企業経営者が所得の一部を隠して盗用できる状況を想定し，最適な資本構成や資産価値を分析した．文献 14) は無限期間の連続時間モデルにおいて，企業の業績情報の開示にコストがかかる状況を想定し，長期投資の下での最適資本構成や企業の債務不履行行動を分析している．文献 2, 3) や文献 4) では有限期間の連続時間モデルで最適契約を分析している．また，モデルの不確実性に関しては，例えば文献 9) 等のロバスト制御（robust control）の文献を挙げることができる．それぞれ連続時間での確率解析の利便性を利用して，情報問題の分析に関して従来の期待効用アプローチの枠組みを超えた新しい経済分析の分野を切り開いている．詳しくは引用した文献等を参照してほしい．

一方，後者の効用関数の性質に関しても，最近になって連続時間での確率解析の手法が数多く利用され経済分析が大きく発展している．本稿では以下において，その動向を少し詳しく紹介する．

2. 確率微分効用 (stochastic differential utility)

期待効用アプローチでは効用関数の性質に関して，危険回避（risk aversion）と異時点間の代替の弾力性（intertemporal elasticity of substitution, 以下では略して IES）という理論的には全く別の 2 つの概念が，モデル内では表裏一体に特徴付けられてしまうという問題が生じる．もう少し具体的に説明すると，危険回避度 δ とは，ある時点で消費量が不確実な場合に，経済主体が状態間の消費財の相対消費量を価格弾力的にどれだけ変化させたくないか（すなわち状態間で消費量をどれだけ平準化（smoothing）したいか）を表す．つまり危険回避度とは，ある固定した将来時点において経済主体が消費量の不確実性をどれだけ回避したいのかを示す．一方，IES とは，経済主体が異時点間の消費財の相対消費量を価格弾力的にどれだけ許容するか（すなわち異時点間で消費量をどれだけ平準化（smoothing）させたくないか）を示す．つまり IES とは，経済主体が異時点間で消費量の柔軟な変化をどれだけ好むのかを表す．しかし，上述したように，期待効用アプローチでは効用は状態分離可能である上に時間分離可能であるために，時間と状態の識別に関して，その価格弾力性は無差別（indifferent）である．その結果，異時点間で消費を平準化する選好と，ある時点を固定したときに状態方向に消費を平準化する選好が一致する．すなわち IES は危険回避度の逆数 $\frac{1}{\delta}$ となる．言い換えれば，経済主体がある時点で消費の不確実性を回避したいと思っているときは，異時点間の消費の代替の弾力性も低下してしまうことを意味する．このような効用関数の性質が現実経済を分析する上で大きな問題であることが実証的に示されている．

その問題を解決するために経済学では近年，期待効用アプローチに代わって，時間に関して分離不可能であるような効用を使用することが多くなっている．それが再帰的効用（recursive utility）と呼ばれ，特に連続時間での再帰的効用を確率微分効用（stochastic differential utility, 以下 SDU）と呼ぶ．ここで SDU を定式化する．各時点 $t \in [0,T]$ において，消費過程 $\boldsymbol{c}$ から

$$dV_t = -f(c_t, V_t)dt + \Sigma_t^\top dB_t, \quad V_T = 0. \quad (2.1)$$

により特徴付けられる効用 $V_t(\boldsymbol{c})$（あるいは単純に V_t と表す）を得る．なお，肩付き文字 $\top$ は行列あるいはベクトルの転置を表す．ここで，f は集計関数（aggregator）と呼ばれる $\mathbf{R}^2$ から $\mathbf{R}$ への関数である．集計関数 f は，期待効用アプローチのように瞬間的な消費量 c_t に依存するだけでなく，再帰的に効用にも依存しているために時間不可分な効用を表すことができる．ただし，ここでは簡略化のために，f の中に Σ_t は含まれていない．数学的に見ると，この効用は後ろ向き確率微分方程式（backward stochastic differential equations）によって特徴付けられる．ちなみに，経済

学では離散時間モデルでも再帰的効用は分析されているが，その性質は連続時間モデルによってより厳密に一般化できる．

本節では SDU の経済学的な意義を明確に示すために，具体的に文献 15) で導入された集計関数 f の特定の関数形に注目する．すなわち，3 定数パラメータ α, β, δ に対して，

$$f(c,v) \triangleq \begin{cases} (1+\alpha)\left\{\dfrac{c^{1-\delta}}{1-\delta}|v|^{\frac{\alpha}{1+\alpha}} - \beta v\right\} \\ (\delta \neq 1 \text{ のとき}), \\ (1+\alpha v)\left\{\log c - \dfrac{\beta}{\alpha}\log(1+\alpha v)\right\} \\ (\delta = 1 \text{ のとき}) \end{cases} \tag{2.2}$$

という関数型を考える．ここで，β は時間選好を，α は危険回避度を，$\frac{1}{\delta}$ は IES を表す．危険回避度と IES は独立なパラメータとして特徴付けられている．それらの経済学的解釈は以下ですぐに説明する．この集計関数を Schroder-Skiadas（以下，SS）集計関数と呼ぶことにする．ちなみに，過去の SDU の文献では，SS 集計関数の代わりにしばしば Duffie-Eptein（以下 DE）型集計関数が使われる[7]．ここではモデル上の便宜から SS 集計関数を使う．SS 集計関数と DE 集計関数の比較については，例えば文献 13) を参照．

ここで，V_t の存在，V_0 の消費計画 $\boldsymbol{c}$ に対する単調増加性，凹性を保証するためにパラメータに制約を課す．

$$\beta \geq 0 \text{ かつ } \begin{cases} \alpha > -1 \quad \text{かつ} \\ 1-\delta < \min\left\{1, \dfrac{1}{1+\alpha}\right\} \\ (\delta \neq 1 \text{ のとき}), \\ \alpha \leq \beta \quad (\delta = 1 \text{ のとき}). \end{cases} \tag{2.3}$$

パラメータ α, δ の関係を直観的に解釈するために，効用過程 V_t を次のように単調変換する．

$$U_t = \begin{cases} (V_t)^{\frac{1}{1+\alpha}} \quad (\delta < 1 \text{ のとき}), \\ -|V_t|^{\frac{1}{1+\alpha}} \quad (\delta > 1 \text{ のとき}), \\ \dfrac{1}{\alpha}\log(1+\alpha V_t) \quad (\delta = 0 \text{ のとき}). \end{cases} \tag{2.4}$$

SS 集計関数の下では，この V_t の序数的に同値な効用過程 U_t は

$$U_t = \begin{cases} E_t\left[\displaystyle\int_t^T e^{-\beta(s-t)} \times \right. \\ \left. \quad \left(\dfrac{c_s^{1-\delta}}{1-\delta} + \dfrac{\alpha}{2}U(s)^{-1}\,||\sigma_U(s)||^2\right) ds\right] \\ (\delta \neq 1 \text{ のとき}), \\ E_t\left[\displaystyle\int_t^T e^{-\beta(s-t)} \times \right. \\ \left. \quad \left(\log c_s + \dfrac{\alpha}{2}\,||\sigma_U(s)||^2\right) ds\right] \\ (\delta = 1 \text{ のとき}) \end{cases} \tag{2.5}$$

と表される．ここで，上述の期待効用アプローチでの時間分割型効用の定式化を思い出せば，式 (2.5) が，時間分割型効用の部分と確率微分効用に特有の追加的な効用の部分の 2 つに分解できることがわかる．項 $||\sigma_U(s)||^2$ は将来の効用の不確実性を表しており，追加的な効用の部分全体では $\delta < 1$（すなわちすべての $t < T$ について $U_t > 0$）の場合には，$\alpha < 0$ であれば負の効用をもたらし，$\alpha > 0$ であれば正の効用をもたらす．そこで $\delta < 1$ のとき，$-\alpha > 0$ であれば個人は**時間分割型効用に比較してよりリスク回避的**であると言い，$-\alpha < 0$ であれば個人は**時間分割型効用に比較してよりリスク愛好的**であると言う．$\delta = 1$ のときには，U_t の符号は正負両方取り得るが追加的な効用の部分が U_t の符号に依存しないため，α の符号の意味は $\delta < 1$ の場合と同様である．$\delta > 1$ のときには，すべての $t < T$ について $U_t < 0$ となるため，α の符号の意味は逆になる．$\alpha = 0$ であれば，個人は期待効用アプローチのように時間分割型効用を持ち，将来の効用の不確実性に関して無差別である．

3. SDU を使った金利の期間構造分析

次に，経済分析で SDU を使用することの具体的な効果を見てみる．あらかじめその効果を一言で言い表せば，期待効用アプローチに比べ長期の不確実性を追加的に価値評価することである．この効果を具体的に観察するために，文献 14) を参考に，金利の期間構造分析の事例を眺める．

具体的な分析を見る前に，まずは金利の期間構造の，現実経済の中での役割とその分析の意義を簡潔に復習する．いま，満期において 1 単位の消費財の支払いを保証する債券を考える．これを（実質）ゼロクーポン債と呼ぶ．そして $P(t,s)$ は，満期 s のゼロクーポン

債の時刻 t における価格を表すとする．このとき

$$R(t,s) := -\frac{1}{s-t}\log P(t,s)$$

と定義すれば，$R(t,s)$ は時刻 t から s のスポットイールドと呼ばれ，ゼロクーポン債の時刻 t から満期 s までの複利ベースの最終利回りを表す．ある時刻 t を固定し満期までの時間 $s-t$ を横軸にとってスポットイールドをプロットした曲線をイールドカーブと呼び，金利の期間構造を表す．マクロ経済学的には，金融政策を運営する中央銀行やポートフォリオを組成する投資家に対し将来の金利に関する市場情報を与えている．また，金融実務の場においては，莫大な量の債券や金利デリバティブが市場で取引されており，それらの金利関連商品のみならず，ほぼすべての証券の価格付けに金利の期間構造が用いられている．しかし，このような重要性にも関わらず，金利の期間構造と，GDP・物価上昇率などの経済の基本変数や個人の選好との間にどのような構造的な関係があるかはあまり明らかにされていない．例えば，過去のデータ分析から，平均的に見て名目イールドカーブは右上がりであることが知られている[10)]．一方，文献 6) 等で期待効用アプローチを使った金利の期間構造モデルでは，右上がりの実質イールドカーブが現れるのは GDP の成長率が実質短期金利と負の相関を持つ場合に限られるという特徴がある．しかし実証的には GDP の成長率は名目短期金利と正に相関していることが示されている．言い換えれば，景気が良いときには平均的に見れば金利は上がり，景気が悪いときにはその逆である．もし実質金利が名目金利と並行に変動しているとすれば，実質イールドカーブの傾きは右下がりとなる．このように名目ベースと実質ベースの間で金利の期間構造の傾きを整合的に説明することは経済学の大きな課題の一つである．SDU を使った金利期間構造分析はこの課題を解決する一つの試みと位置付けることができる．

具体的なモデルをつかって実際の分析を観察する．この経済モデルには代表的個人がいて，一種類の非耐久的な消費財が存在する．この実物経済 (real economy) には 2 つの状態変数が存在する．第一の状態変数 e は所得である．所得の確率過程は外生的であり，次の確率微分方程式に従う．

$$\frac{de_t}{e_t} = \mu_e(t)dt + \sigma_e^\top dB_t, \quad e_0 \in \mathbf{R}_+. \quad (3.1)$$

ここで $\mu_e(t) \triangleq \nu_t - \beta$ であり，また $\sigma_e \in \mathbf{R}^{2\times 1}$ は定数ベクトルである．第二の状態変数は所得の期待成長率 ν であり，次の確率微分方程式に従う．

$$d\nu_t = k(\bar{\nu} - \nu_t)dt + \sigma_\nu^\top dB_t, \quad \nu_0 \in \mathbf{R}. \quad (3.2)$$

$\sigma_\nu \in \mathbf{R}^{2\times 1}$ は定数ベクトル，$\bar{\nu}$ は定数である．$\bar{\nu}$ は所得の期待成長率の平均回帰のレベル，k は平均回帰のスピードを表すパラメータである．代表的個人は消費財を消費することで上述の SDU に基づいて効用を得る．

式 (2.3) の下で，均衡における効用過程 V_t は次の確率微分方程式の解として特徴付けられる．

$$\begin{cases} \dfrac{de_t}{e_t} = \mu_e(t)dt + \sigma_e^\top dB_t, \quad e_0 \in \mathbf{R}_+, \\ d\nu_t = k(\bar{\nu} - \nu_t)dt + \sigma_\nu^\top dB_t, \quad \nu_0 \in \mathbf{R}, \\ dV_t = -f(e_t, V_t)dt + \sigma_V(t)^\top dB_t, \quad V_T = 0. \end{cases} \quad (3.3)$$

ここで $t \in [0,T]$ において $V_t = J(t, e_t, \nu_t)$ を満たす $J \in C^{1,2}([0,T]\times\mathbf{R}^2)$ が存在すると仮定する．このとき，効用過程 $V_t = J(t, X_t)$ は $\sigma_J \triangleq e\sigma_e\frac{\partial J}{\partial e} + \sigma_\nu\frac{\partial J}{\partial \nu}$ として

$$dJ_t = -f(e_t, J_t)dt + \sigma_J(t)^\top dB_t \quad (3.4)$$

を満たすことがわかる．

SS 集計関数の下では，満期 s のゼロクーポン債の時刻 t における価格 $P(t,s)$ は，

$$P(t,s) = E_t\left[\exp\left\{-\int_t^s r_u du\right\}\times \exp\left\{-\int_t^s \lambda_u^\top dB_u - \frac{1}{2}\int_t^s ||\lambda_u||^2 du\right\}\right]$$

と表され，したがって，時刻 t から s のスポットイールド $R(t,s)$ は図 1 のようになる．特に，$\delta = 1$ のケースについては，均衡スポットイールド $R(t,s)$ を解析的に閉形式で以下のように解くことができる（詳しくは文献 13) を参照）．

$$\begin{aligned} &R(t,s) \\ &= r_t + (\bar{\nu} - \nu_t)\left(1 - \frac{1-e^{-k(s-t)}}{k(s-t)}\right) \\ &- \frac{\sigma_\nu^\top\sigma_e}{k}\left(1 - \frac{1-e^{-k(s-t)}}{k(s-t)}\right) \\ &- \frac{||\sigma_\nu||^2}{2k^2}\left(1 - 2\frac{1-e^{-k(s-t)}}{k(s-t)} + \frac{1-e^{-2k(s-t)}}{2k(s-t)}\right) \\ &+ \underbrace{\frac{1}{k}\frac{\alpha}{\beta}\left(\sigma_\nu^\top\sigma_e + \frac{||\sigma_\nu||^2}{k+\beta}\right)\left(1 - \frac{1-e^{-k(s-t)}}{k(s-t)}\right)}_{\text{追加的な期待所得ショック}}. \end{aligned}$$

この式を使って，上述した期待効用アプローチの問題点を改めて振り返りたい．まず，文献 13) に示されているように，均衡では σ_ν は瞬間的スポットレート r_t の分散項を表す．したがって，第 3 項の $\sigma_\nu^\top\sigma_e$ は，実質短期金利と景気循環の相関を表している．ここで上

$$
R(t,s)=\begin{cases}
-\dfrac{1}{s-t}\log E_t\left[\begin{array}{c}\exp\left\{-\displaystyle\int_t^s\left(\begin{array}{c}\beta+\delta\mu_e(u)-\dfrac{\delta}{2}||\sigma_e||^2\\+\dfrac{\alpha(1+\alpha)}{2}\left|\left|\dfrac{\sigma_J(u)}{(1+\alpha)J_t}\right|\right|^2\end{array}\right)du\right\}\times\\ \exp\left\{-\displaystyle\int_t^s\left(\delta\sigma_e-\alpha\dfrac{\sigma_J(u)}{(1+\alpha)J_t}\right)^\top dB_u\right\}\end{array}\right] & (\delta\neq 1\text{ のとき}),\\
-\dfrac{1}{s-t}\log E_t\left[\begin{array}{c}\exp\left\{-\displaystyle\int_t^s\left(\begin{array}{c}\beta+\mu_e(u)-\dfrac{1}{2}||\sigma_e||^2\\+\dfrac{\alpha^2}{2}\left|\left|\dfrac{\sigma_J(u)}{(1+\alpha J_t)}\right|\right|^2\end{array}\right)du\right\}\times\\ \exp\left\{-\displaystyle\int_t^s\left(\sigma_e-\alpha\dfrac{\sigma_J(u)}{(1+\alpha J_t)}\right)^\top dB_u\right\}\end{array}\right] & (\delta=1\text{ のとき}).
\end{cases}
$$

図 1　スポットイールド $R(t,s)$.

述の議論を参考に実質短期金利は景気変動と正の相関を持っていると仮定する．すなわち第 3 項はマイナスである．しかも第 4 項は，時間的割引要因を表す指数関数による凸性効果を表すので，傾きを押し下げる効果をもたらす．もし第 2 項をゼロ（期待所得成長率が $\bar{\nu}$ に等しい）とすれば，期待効用アプローチでは第 5 項は存在しないこともあって，均衡実質スポットイールド $R(t,s)$ のカーブは必ずマイナスの傾きを持つ．特に第 4 項の凸性効果は満期までの長さの 2 乗のオーダーで強く実質イールドカーブの傾きを押し下げるため，これを前提に名目イールドカーブがどのように右上がりの傾きを形成するのかは，期待効用アプローチを使ったときの問題点である．それに対し，SDU では $\alpha\neq 0$ なので，将来の効用の不確実性を評価することから追加的に第 5 項が存在し，長期の不確実性を追加的に評価できる．もちろん実際の数値例では第 5 項だけで傾きが反転するほど大幅に問題が解決するわけではは必ずしもない．しかし，金利の期間構造をモデル化する際に，期待効用アプローチとは対照的に危険回避度と IES を分離することによって，リーズナブルな形でモデルの自由度が向上している．なお，$\delta\neq 1$ の場合には，一般には J の解析的な閉形式での表現が得られないため，金利の期間構造の閉形式での表現を得ることは難しく，数値解析的に解くこととなる．

ここまでは，物価変動を考慮せず実物経済のみをモデル化してきた．上述した実物ゼロクーポン債は実物経済で満期において 1 単位の消費財の支払いを保証するものであった．しかし，現実においては多くの場合，債券のクーポンや元本等の支払い額は消費財単位でなく貨幣単位で定められ，その結果物価変動リスクを持っている．そこで物価指数をモデルに導入することで名目イールドカーブを導出する．すなわち，状態変数として物価指数 N とその期待成長率（期待物価上昇率）ε の 2 つを追加する．なお，ここまでブラウン運動を 2 次元としていたが，上の変更にあわせてこれまでに定められた確率過程も適宜修正する．2 つの変数の従う確率過程として，特に期待物価上昇率が平均回帰的になるように以下のように定める．

$$\frac{dN_t}{N_t}=\varepsilon_t dt+\sigma_n^\top dB_t,\quad N_0\in\mathbf{R}_+, \tag{3.5}$$

$$d\varepsilon_t=\theta(\bar{\varepsilon}-\varepsilon_t)dt+\sigma_\varepsilon^\top dB_t,\quad \varepsilon_0\in\mathbf{R}. \tag{3.6}$$

ここで，θ（平均回帰のスピード）は正の定数，$\bar{\varepsilon}$（平均回帰の水準）は定数，$\sigma_n,\sigma_\varepsilon\in\mathbf{R}^4$ は定数ベクトルである．

以上の設定の下で，実質・名目のイールドカーブのマクロ経済学的意味を考察する．特にここでは $\delta=1.5$ の場合に着目する．α は $\alpha\in\{0,\ 3,\ 9\}$ とする．また時間選好のパラメータ β を 0.01 とする．状態変数の分散共分散のパラメータを表 1 のように定める．簡単化のため ρ_{en}, $\rho_{\nu n}=0$ と定める．これにより金利の期間構造に対する期待物価上昇率の影響に焦点を当てることができる．また，経済の終期 T は十分大きく 400 年とし，現時点 t を 0 時点としている．数値解析的には，式 (3.3) を，文献 8) のモンテカルロ法による解法を用いて数値的に解いている（詳細については文献 14) を参照）．数値計算の結果は図 2 のとおりである．

表 1　パラメータの値.

β	0.01	α	0, 3 or 9
$\|\|\sigma_e\|\|$	0.05	$\|\|\sigma_n\|\|$	0.05
ν_0	0.03	ε_0	0.03
$\bar{\nu}$	0.03	$\bar{\varepsilon}$	0.03
k	1	θ	0.05
$\|\|\sigma_\nu\|\|$	0.02	$\|\|\sigma_\varepsilon\|\|$	0.01
$\rho_{e,\nu}$	0.5	$\rho_{e\varepsilon},\rho_{\nu\varepsilon}$	−0.5
		$\rho_{en},\rho_{\nu n},\rho_{\varepsilon n}$	0

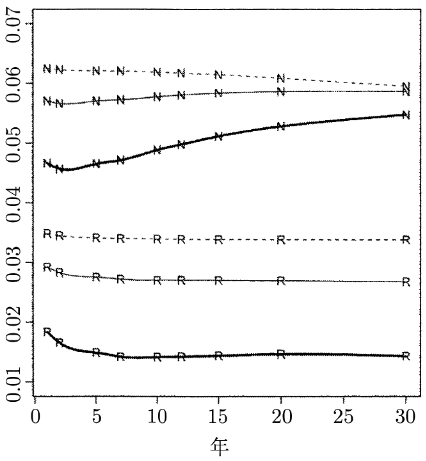


図2 数値計算の結果．N は名目イールド，R は実質イールドカーブを表し，--- $\alpha=0$，— $\alpha=3$，**—** $\alpha=9$ とする．

まず，実質イールドカーブのスロープについては，凸性効果がイールドカーブの傾きを押し下げ，しかも $\sigma_e^\top \sigma_\nu > 0$ の下では，債券は将来の効用の不確実性を軽減することから，個人はリスクヘッジのためにリターンが低い状況でも債券を購入する．さらに，非時間分離型特有の項が個人の長期の債券を購入する動機を増大させ，イールドカーブの傾きを押し下げる．$|\alpha|$ が大きければ，この効果は大きくなる．ただし，この数値例においては，所得の期待成長率の平均回帰のスピードが比較的高いため，スケールの都合上，イールドカーブはほぼフラットに見える．次に，名目イールドカーブを見ると，まず名目の凸性効果は常にスロープを右下がりにする方向で作用する．しかし，$\rho_{e\varepsilon}=\rho_{\nu\varepsilon}=-0.5$ の下では，長期の名目債券を売却する動機が十分に強いので，数値結果ではイールドカーブは右上がりになっている．$|\alpha|$ の値が大きくなるとスロープも急になる．このように，危険回避度と IES を分離することによって，期待効用アプローチに比べ長期の不確実性を追加的に価値評価できるのが SDU の特徴である．

4. おわりに

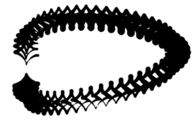
この節では，SDU の議論に絞って連続時間の確率解析の経済学への応用を見てきた．このほかにも第1節で簡単に触れたように，今後連続時間の確率解析の経済学への応用は一層広がりを見せると思われる．

参考文献

1) Black, F., and John Cox (1976): "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, **31**, 351–367.
2) Cvitanić, Jakša, Xuhu Wan, and Jianfeng Zhang (2006): "Optimal Contracts in Continuous-Time Models," *Journal of Applied Mathematics and Stochastic Analysis*, vol.**2006**, Article ID 95203, 1–27.
3) Cvitanić, Jakša, Xuhu Wan, and Jianfeng Zhang (2009): "Continuous-Time Principal-Agent Problems with Hidden Action and Lump-Sum Payment," *Applied Mathematics and Optimization*, **59**, 99–146.
4) Cvitanić, Jakša, and Jianfeng Zhang (2007): "Optimal compensation with adverse selection and dynamic actions," *Mathematics and Financial Economics*, **1**, 21–55.
5) Cox, J., J. Ingersoll and S. Ross (1985a), "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, Vol.**53**, pp.363–384.
6) Cox, J., J. Ingersoll and S. Ross (1985b), "A Theory of the Term Structure of Interest Rates," *Econometrica*, Vol.**53**, pp.385–407.
7) Duffie, D. and L. Epstein (1992), "Stochastic Differential Utility," *Econometrica*, Vol.**60**, pp.353–394.
8) Gobet, E., J.-P. Lemor, and X. Warin (2005), "A Regression-Based Monte Carlo Method to Solve Backward Stochastic Differential Equations," *Annals of Applied Probability*, Vol.**15** (3), pp.2172–2202.
9) Hansen, Lars Peter, Thomas J. Sargent, Gauhar Turmuhambetova, and Noah Williams (2006): "Robust control and model misspecification," *Journal of Economic Theory*, **128**, 45–90.
10) Homer, S. and R. Sylla (2005): "A History of Interest Rates," Wiley.
11) Kreps, David M. (1988): "Notes on the Theory of Choice", Westview Press: Boulder, Colorado.
12) Merton, R.C. (1974): "On the Pricing of Corporate Debt: the Risk Structure of Interest Rates," *Journal of Finance*, **29**, 449–470.
13) Nakamura, H., K. Nakayama, and A. Takahashi (2008), "Term Structure of Interest Rates under Recursive Preferences in Continuous Time," Asia-Pacific Financial Markets, **15** (3–4).
14) Nakamura, H., W. Nozawa, and A. Takahashi (2009), "Macroeconomic Implications of Term Structure of Interest Rates under Stochastic Differential Utility with Non-Unitary IES," Asia-Pacific Financial Markets, **16**(3).
15) Schroder, M., and C. Skiadas (1999), "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," *Journal of Economic Theory*, Vol.**89**, pp.68–126.
16) DeMarzo, P.M., and Y. Sannikov (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, **61**, 2681–2724.

（なかむら・ひさし，東京大学大学院経済学研究科）

伊藤清先生と確率微分方程式



池田　信行

1. 創造の芽を育む日々

伊藤清先生は少年時代，鈴鹿の山並みが望める祖父の家を訪ね，目にする豊かな自然に限りない好奇心を寄せていた．続いて，父上が歴史と漢文を担当しておられた三重県立神戸（カンベ）中学で，物理を学び数式で表される自然の秩序があることを知り，数学への思いを深めた．4年から（旧制）第八高等学校へ進学の折にドイツ語を第一外国語とする理科乙を選んだのは，将来の数学の学習に備えるためである[1)]．

2. 新たな旅立ち

東京帝国大学理学部数学科に進学したときは，夏目漱石の『三四郎』ほどではないが，名古屋と東京の街の気風の違いに戸惑うとともに，同級生たちがすでに数学の研究の話をしていることが驚きであった．その頃先生は無秩序な現象の中に潜む統計的法則に関心を持っていたが，そのような話題は日本では当時，八高で近藤鉦太郎教授に学んだ現代数学の推論の進め方とは違い，P.L. Laplace の時代と大差ない形で扱われていた．こんなことで研究の明確な課題が定まらない状態で卒業を迎えたが，その直後に一年のときに偶然知った A.N. Kolmogorov の名著『確率論の基礎概念』を今一度読み返したときにそれまで抱いていた朦朧としていたものが晴れた．これを契機に，先生はこの本で示された考えにそって確率論の研究を始めた．その頃の成果が2篇の論文として発表されている．

3. 確率微分方程式の誕生

生涯にわたる先生の研究の方向が決まったのは出版された直後の P. Lévy[3)] を知ったときである．この本では1920年代の初め頃，1913年に出版された J. Perrin: *Le Atomes* に強い影響を受けながら，N. Wiener が導入に成功した連続関数の空間上の Wiener 測度を用いて Brown 運動が論じられている．先生はこの本[3)]で，八高のとき以来色々な話題の学習を続けながら明快には理解できなかった，偶然性を伴う運動がすでに数学の課題として確立されていることを知った．早速この本に取り組み，読み進みながら研究を始めた．その最初の成果が1942年の先生の論文：On Stochastic Processes (I)（*Japan J. Math.* **13** (1942)），である．実は当初これに続く (II) が用意されていた．ところが，太平洋戦争を目前にして，本格的な数学の論文を発表する場はなく，(II) はそのときは日の目を見なかった（1951年に別の形で発表されている）．そこで予定されていた内容の一部が日本語で書かれ，「全国紙上数学談話会誌」の1077号（1942）に発表されている．

Brown 運動を特徴付ける特性の中から，軌跡が連続と言う仮定だけを除いた運動は加法過程と呼ばれる．その軌跡の構造が解析され，軌跡の表現が先に述べた先生の論文 (I) で示されている．これは Lévy が先駆者らしい彼特有の推論で行ったものを，J. Doob の結果の助けを借りて，Kolmogorov 流の厳密な表現に書き換えたものである．ここで得られた軌跡の表示は今日 Lévy–伊藤分解と呼ばれていて，無限分解可能な確率の Fourier 変換に関する Lévy–Khinchin の標準形に対応している．この関係は後で述べる Perrin が指摘している考え方の一例で，D. Mumford も「3000年へ向けての数学」と題する会合で取り上げている．

Perrin の精密な実験により，1ミクロン（$=10^{-6}$ m）位の大きさの粒子が水中（空中）に浮遊するときに Brown 運動が典型的な形で見られことが知られている．類似な運動は顕微鏡が発明された頃 A. von Leeuwenhock により観察され，1827年には R. Brown が花粉の中に含まれる微粒子や London 上空の塵等多種多様な微粒子を観測し同様な運動を確認している．それどころか，紀元前60年に出版された T.C. Lucretius の『物の本質について』（第二巻，項目114-141）で，扉や

鎧戸の隙間を通って薄暗い部屋の中に射しこんでくる太陽の光の中で微粒子が乱舞している様子が語られている．これは，L. Pasteur が指摘しているように，部屋の中に浮遊する塵の中には 2〜3 ミクロンくらいの大きさの微粒子がたくさん含まれていて，Lucretius はそれらの動きを見ている．この室内の塵の話は最近のダスキンのテレビ・コマーシャルで取り上げられている．このような微粒子の不規則な運動は，液体や気体の分子が微粒子に四方八方から絶え間なく衝突するために起こる．同様なことは，黄沙の例のように液体や気体自身が動いていたり，それらの物理的な特性が場所ごとに違う場合にも起こる．そのような場合もこめて，乱舞する微粒子の運動を抽象化して，数学の概念として捉えたのが Kolmogorov[2] の「確率的力学系」である．J.C. Maxwell が 19 才のとき指摘しているようにこのような運動の考察には確率の考えが必要である．一方これらの法則の解明には積分方程式や偏微分方程式等の解析的方法が古くから用いられている．Perrin はノーベル賞受賞講演で，一つのものを表，裏の両面から見る二つの方法による考察の重要性を指摘している．Kolmogorov は確率的力学系をある種の積分方程式をみたす推移確率の族を用いて捉えた．これに対して，伊藤先生は Lévy が加法過程について行った眺め方にならって，加法過程の軌跡を接線とする運動方程式を用いて確率的力学系の解明を目指した．これが，確率微分方程式の話の始まりである．このために通常の力学の場合にならい，目標とする運動方程式を積分方程式として捉える考えが浮ぶ．しかし話を軌跡が連続の場合に限定しても，Brown 運動の軌跡は Wiener が示したように，(大胆な言い方をすれば) 次数が 1/2 の Hölder 連続な関数であるために新たな障害が現れる．先生は「Brown 運動による確率積分」の概念を導入し，この困難を解消した．この議論では，Brown 運動が加法過程の特性を持つことと，任意の時間区間における軌跡の位置の変化の増分の 2 乗の平均が区間の長さになることが大きな役割を果たしている．この確率積分の概念とそれに関連する伊藤の公式が，その後の理論の発展で基本的な役割を担っている．これらについては多くの所で語られているのでここではこれ以上立ち入らない．なお，先生に始まる理論の発展では，Doob が指摘した「理論をマルチンゲールの概念を基礎にして捉える」考えが柱の役割を担っている．

4. ただよう暗い影

確率微分方程式は大変困難な社会的環境の中で生まれたが，その頃日本の外でも大きい悲劇が起きた．先生と同年齢の W. Doeblin は Lévy の本[3] に心惹かれ，確率的力学系の軌跡に関心を寄せていたが，ナチとの戦いで 1940 年 6 月 20 日ドイツ軍に包囲され，自ら短い生涯を閉じた．先生より 2 才若い K.L. Chung（鐘開来）は天津大学の図書が日本軍の爆撃を避けるために昆明（Kunming）の地下庫倉に疎開されていたので Lévy の本[3] を読む機会を奪われた無念さを生涯忘れなかった．1940 年頃までの Lévy の研究資料はナチとの戦いで焼失した．困難は戦後も続き，Cho Ju-Gyoung（ジョ・ジュギョン）（金日成大学教授）は，確率微分方程式について体系的に論じた世界で最初の単行本を 1962 年に出版し，伊藤先生に送ってきたが先生と会う機会を持てないままに他界した．

5. むすび

伊藤先生は学習院大学を定年後，若い頃試みながら踏み込むことができなかった，基本になる方程式の立て方から始まる数理物理の研究を目指していた[1]．安野光雅:『スイスの谷』に，福沢諭吉（『文明論之概略』，82–83 頁）が生菓子屋の仕入れの話を例にして J. Bernoulli の大数の法則について論じていることが紹介されている．この紀行文の結びは，もう一人のスイス生れの偉大な数学者 L. Euler と Bernoulli の 2 人とスイスの文化や産業の交わりの話である．外国の旅人が伊勢地方を訪ね，第一回 Gauss 賞受賞者伊藤清先生を偲び，日本の伝統と数学のかかわりを語る日が来ることもまんざら空想ではないかもしれません．

参考文献

1) 伊藤清，「数理物理への回帰」学術月報 Vol.**43**，(1990 年)．その他，先生が書かれた i)「数学の研究を始めた頃」，昭和一人一話集，道，ii) 近藤鉦太郎教授と数学，近藤鉦太郎教授記念論文集（八高）や，iii) 傘寿の際の先生の講演記録等を参考にした．
2) A.N. Kolmogorov: Über die analytischen Methoden in der Wahrscheinlichkeitsrenung, *Math. Ann.*, **104** (1931).
3) P. Lévy: "Théorie de l'addition des variables aléatoires" (1937).

（いけだ・のぶゆき，大阪大学名誉教授）

確率論のおもしろさ

楠岡　成雄（聞き手：編集部）

1. 確率論は今どうなっているか

——確率論は様々な分野で使われているようですが，確率論そのものの現在の研究状況などをお聞かせいただけますでしょうか．

確率論では様々な研究がありますが，歴史的に見ると大きな流れとして，1960年代にマルコフ過程論，70年代に確率微分方程式，80年代にマリアバン解析の研究が盛り上がりました．その後90年代にフラクタル上の確率過程論が流行しましたが，だんだん大きな流行がなくなってきたように思います．



楠岡成雄氏．

伊藤先生の確率微分方程式がよい例ですが，確率論では，節目節目に新しいアイディアが現れ，それが大きな理論となり，確率論全分野における基本的な道具となることがこれまでありました．確率論の分野でパーコレーションに関する研究により2006年にヴェルナーがフィールズ賞を取りました．ヴェルナーらの業績は，数学の観点から見れば，2次元特有の理論になっているので，それがそのまま確率論の基本的道具になるようには私には思えません．確率論では現在，様々な分野で昔から問題にされていることの研究が少しずつ進展し深まっていますが，確率論の全分野に影響を与えるような研究はなかなか出てこないのではないかという印象を持っています．

例えば，確率論では古くからランダムウォークの話が扱われています．普通の形のグラフにおける理論はかなり完成しているのに対して，不規則なグラフの上でのランダムウォークの性質はあまりよく分かっていませんでした．それがフラクタル上の拡散過程の研究あたりから少しずつ進展しています．これは非常に古典的な問題ですが，今まで全く手が付かなかったところに手が付き始めました．しかし，同時に研究内容は極めて複雑になりつつあります．

かつて京都大学の熊谷隆君がフラクタル上の拡散過程に関するユニークな論文を学生のときに書

きました．当時は確率積分などの既存の道具が一切使えなかったので，アイディアで勝負するしかなかったわけです．そのため修士1年生でもそのような論文を書けたのですね．確率論には，問題をしっかり理解してアイディアがあれば，あまり確率論の様々な知識がなくても，解けてしまう問題が少なからずありました．今ではそういった問題がなかなかなくて，相当勉強しないと研究の最前線に到達できないような世界になってしまいました．色々ある問題のうち，簡単なものはどんどん解かれてしまって，難しいものが残ってしまったわけです．かつてのフラクタルのような新しい研究対象が現れれば，若い研究者の活躍の場ができてよいのですが，難しいようです．

確率論の問題とは何かということですが，どんなものからでも確率論の問題が作れます．整数論も対象となります．もし素数がランダムだとしたらどういうことが予想できるかということを実際に証明したりしています．あるいは多項式の係数がランダムだったら解がどういう分布をするかということも研究対象になります．このように，確率論が研究対象とするものは大変広く，その中には物理学・生物学などの実際の現象に対するモデルからきた問題もあれば，今述べたような全くの数学的好奇心からきた問題もあります．

確率論には，かつてのフェルマー予想のように，単純に定式化でき，大学1年生くらいならば問題の意味が理解できるが長い間未解決となっている問題もあります．例えば自己回避確率過程に関する予想というものがそうです．このような問題は統計物理に多く，物理学者はこうなるはずだ，と答えを予想しているけれど数学的には証明できないものがたくさんあります．問題を解析する数学的道具がないため，このような問題に手が付けられません．何か深い構造が見つかれば研究が進展するのですが．

統計学との関連も重要です．物理学ではモデルに現れる定数をすべて測定で決めることができますが，例えば生物学では突然変異がどのくらいの頻度で起こるかというのは簡単には分かりません．過去の様々なデータを見て推定するわけですが，突然変異の起こる頻度がある時点で変わっていたら，定数を定めるのは単純にはできません．そこで，ある種の仮説をおき，それをもとに複雑なモデルをたてて，現実と照らし合わせるわけです．これは統計学の問題となります．統計学は確率論と表裏一体ですから，統計学への応用をにらんだ確率論の研究は大事です．こういった方向の研究は盛んになされていますが，何か新しいことが出てくるのではないかと期待しています．

2. 社会への浸透

――現在では，確率論はどのような所に応用されているでしょうか．

そうですね．物理学や生物学に限らず，化学，経済学，工学と様々に応用されているように思います．

ノーベル賞を受賞したフランスの物理学者ピエール＝ジル・ド・ジェンヌ氏が高分子の本を書いていますが，その中では確率論の考え方が用いられていて，多くの確率論の命題が述べられています．物理学では多くの場合，数学的に厳密な議論はなされないので，物理学は確率論研究者にとり未解決問題の宝庫です．

ところで，物理学の方程式はすべて可逆的です．ニュートンの方程式を見ても，初期条件の位置はすべて同じで，速度を全部マイナスにすればちょうど時間反転します．そうすると，すべての物理現象というのは，可逆的なわけですが，我々の見ている世界はほとんど可逆的ではありません．統計力学では様々な現象を確率論的に説明するわけですが，そういうものを古典的な物理法則から説明，導出ができるだろうかということが問題になります．確率論ではそういったことの研究も行われています．

生物学は非常に範囲が広いのですが，特に確率論が盛んに使われているのは，集団遺伝学のような遺伝の話ではないかと思います．木村資生先生が遺伝子レベルでの突然変異について中立仮説を唱えられましたが，その後，遺伝に関しては様々な確率モデルがたてられ研究されています．また，生態学でも確率モデルが使われているようです．

化学のことはよく知りませんが，伊藤清先生が京都賞を受賞されたとき，同じく高分子化学分野で京都賞を受賞されたクルト・ヴュートリッヒ氏が確率微分方程式を使って論文を書いたことがあると受賞式でおっしゃっておられました．

経済学は，実験ができないせいか，数学的に厳密な議論をきっちり行うようです．確率モデルを用いたり，理想状態を作り出すために確率論的な発想を加えたりしています．ただ，数学的には難しくなることも多いようです．最近は数値計算に頼ることもかなりあると聞いています．

工学でもモンテカルロ法という手法があります．これは人為的にランダムな数列，乱数を作り出し，それを利用して問題を解こうとする手法ですが，確率論により方法が正当化されます．

このように，確率論の知識，考え方はありとあらゆる分野に非常に浸透しており，利用されていると思います．数学としての確率論の研究は何であっても，それらの応用と繋がっていると思います．

3. 確率論を学ぶための心得

――様々な現状の中で，確率論を修得していくためにはどのような道を通る必要があるでしょうか．

そうですね．ご存じのように確率論は測度論を数学的な基盤として構築されていて，様々な道具を使います．測度論をちゃんと理解しようとすると，非常に抽象的な数学の知識が必要になり，その中でも集合と位相をきっちり勉強する必要が出てきます．確率論を専門としたい人にはこういったことをしっかりと勉強してほしいと思いますが，現在では様々な参考書も出ていますから，他分野の人が最初に勉強するときには，あまり数学的な議論は気にせず，まずは自分の研究している分野に密着したことから学習するのがよいのではないかと思います．

確率論の基礎というのは非常に単純なものです．しかし応用分野に応じて多種多様な理論ができあがっているので，どの理論を使うかにより勉強すべきことが変わってきます．ただ，本格的に使えるようになるためには最終的には測度論を勉強してほしいと思います．

確率論の展開においては様々な数学的な知識を用いますので，確率論の研究者を志す人は，数学全般にわたって様々なことを知っていたほうがいいように思います．

4. 影響を受けたこと

――先生が確率論を学び，研究されてきた中で影響を受けたことなどをお聞かせいただけますでしょうか．

そうですね．やはり，マリアバン解析でしょうか．

マリアバン解析は，ノバート・ウィナーが提唱したウィナーカオスと呼ばれるものあたりが出発点となります．日本でも，伊藤先生や丸山儀四郎先生がウィナーカオスを非常に意識しておられました．

マリアバン解析以前の確率論の主な解析手法は，基本的には確率微分方程式の考え方です．マリアバン解析では，確率微分方程式の解をウィナー以来の考え方に基づき，ブラウン運動の汎関数として捉えます．今ではブラウン運動の汎関数はウィナー汎関数と呼ばれるようになりました．基本的には，ブラウン運動の道筋が一本与えられると，それに対応して確率微分方程式の道筋が一本決まるというような考え方です．この視点で確率微分方程式を扱うには，独自の新しい解析を作ればよいというのがマリアバンのアイディアです．

マリアバンの論文の出たすぐ後に，池田信行先生と渡辺信三先生がマリアバンのアイディアをいかに厳密に実現させるかということを研究し始められました．そして，マリアバン解析で色々なことができるということが分かりました．そのことが，日本においてマリアバン解析が非常に早い段階で広まった理由です．同じ頃ストルックやビスミュといった欧米の研究者が，別の定式化を考えるという形で研究が進展していきました．

私はストルック氏と1980年代に一緒に研究することになりました．彼とは歳が離れていますが非常にウマが合い，未開の分野だったので，最初のところから延々と議論をして積み上げていきました．未発表の結果や当時解くことができなかった問題もたくさんあります．ここ数年で，ファイナンスにおける数値計算の研究（計算ファイナンスと呼ばれています）にマリアバン解析が使えることが分かってきました．その流れで，当時未解決だった問題を最近解くことができました．

私がファイナンスの研究を始めた20年近く前ですが，そのときはマリアバン解析が実際的にファイナンスに使えるとは誰も思っていませんでした．数学というものは全く関係がないように見えても繋がっているようです．また，そのとき研究していることと全く違う分野のアイディアを取り入れると問題解決に繋がることがあります．

数値計算・数値解析は数学における研究分野として今後大変重要になってくると思います．数学モデルが複雑になり，数値計算を用いないと答えが出せなくなりつつあります．数値解析には単純な問題も山のようにあります．逆行列を求める問題でも全面的に解決されているわけではありません．あらゆる数学の問題に対して数値解析の問題が存在するし，様々な数学の分野のアイディアが数値解析の問題解決に使えるだろうというふうに私は思っています．そういう意味では数値解析の分野に数学者はもっと興味をもっていいと思います．

5. メッセージ

――学生や後進の方たちに何かメッセージがありましたらお願いします．

確率論は，数学の理論でありますが，同時に物理学を初めとした他の様々な科学分野の問題とも絡んでいます．

確率微分方程式やマリアバン解析などは，わりと数学の内在的な動機から出発したものという印象を持っていますが，その発想においては，物理学などの諸科学からの影響を受けています．

やはり，数学の様々な分野や数学以外のことにも目を向け，好奇心をもって見ることが大事だと思います．そういうことは，確率論を研究するときに限らずいつでも大事なことだと思います．

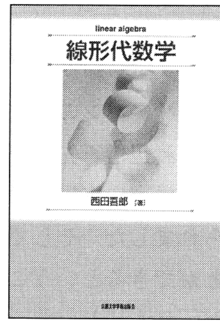
――本日は貴重なお話をお聞かせ下さり，誠にありがとうございました．

(2009年10月，東京大学駒場キャンパス研究室にて)

（くすおか・しげお，東京大学大学院数理科学研究科）

線形代数学

西田吾郎著，A5 判，280 頁，本体 2500 円，京都大学学術出版会



昨年新しく出た本書は，その「はじめに」にあるように，大学の 1 年次に理系の学生が学ぶ内容を主としていて，講義における教科書，もしくは独習書として書かれている．全体の構成としては，第 1 章から第 5 章までで，1 年間の授業でカバーすべき「線形（線型）代数学」の基本がおさめられており，第 6 章，第 7 章がより専門的な内容である．

それに 43 頁分の「付録」があり，高校における数学の補足とともに，多項式とその根，対称式，代数学の基本定理，などが述べられているが，これは本論に対する予備知識を補う意味も持っている．

全体の分量の配分もよく考えられている．それとともに「行列と行列式を用いた計算技術」と「ベクトル空間の理論的解明」とのダイナミックな関わりが強調されている．

各章の本文中にも問題がちりばめられており，章末の問題も充実している．本の最後にはこれらの問題の「解答とヒント」がかなり詳しく書いてあって 27 頁分ある．

まず各章の目次を見てみよう．

第 1 章　ベクトル空間
第 2 章　行列
第 3 章　行列式
第 4 章　線形変換
第 5 章　計量ベクトル空間
第 6 章　行列の指数関数
第 7 章　テンソル積と外積ベクトル空間

第 1 章では，ベクトル空間の理論が述べられる．ここは高校の数学とは雰囲気ががらりと変わり，理論的訓練の素地を作るところなので，講義では教える側，教わる側，双方に忍耐が必要である．教える側としては，学生がここでつまずかぬように丁寧に講義する必要がある．このテキストではいろいろなベクトル空間の具体例が与えてあり，工夫されている．

なお，TEX での記号として，部分集合 S で張られる部分空間は $< S >$ よりも ¥langle, ¥rangle を用いて $\langle S \rangle$ とする方が望ましい．

第 2 章「行列」では，計算法としては「掃き出し法」が重要であり，行列の階数，逆行列，連立一次方程式の解，などの計算に使われる．学生にとってはパソコンでの表計算などをイメージすればなじみ易いであろう．

第 3 章「行列式」において，**行列式**の定義は，天下り式に与えるのではなく，n 次正方行列 A を $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ と数ベクトルを用いて書いたとき，$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ の n 重交代写像として定義する．ここでは全体として理論が多いので，行列式の計算の反復練習を演習などで補う必要がある．独習者は簡単な演習問題のテキストを用いて計算技術を会得するのがよい．

また，行列式はつねに $\det A$ と書かれ，通常の記号 $|A|$ は避けられているが，通常の記号にも演習などで慣れておくのが望ましい．

さらに本章では，2 つの多項式 $f(x), g(x)$ が共通根を持つための条件を書く**終結式**や，$f(x)$ が重根を持つための条件を書く**判別式**が与えられている．終結式は和算家「関孝和」が『**解伏題之法**』（重訂 1683）で，また判別式は『**開方翻變之法**』（1685）で，それぞれ世界で初めて導入した．これらは行列式の起源であり，このことについて，折角なので是非一言コメントが欲しい（参照 http://hdl.handle.net/2433/25865）．

第 4 章「線形変換」では，固有値・固有ベクトルを中心として述べられている点に特色がある．行列のジョルダン標準形と関連する事柄についても詳しく述べられている．

第 5 章「計量ベクトル空間」では，実数 $\boldsymbol{R}$ 上の場合と，複素数 $\boldsymbol{C}$ 上の場合とが，並行してコンパクトにまとめて述べられている．全体として内容が多いので，講義では適当に取捨選択する必要がある．

しかしながら，内容豊富という点では，「講義が終わった 2 年次以降でもこの本を持っていればいろいろ参考にできる」ということになり，お買い得である．

以上をまとめて感想を言えば，教科書または参考書として，良書であると言える．

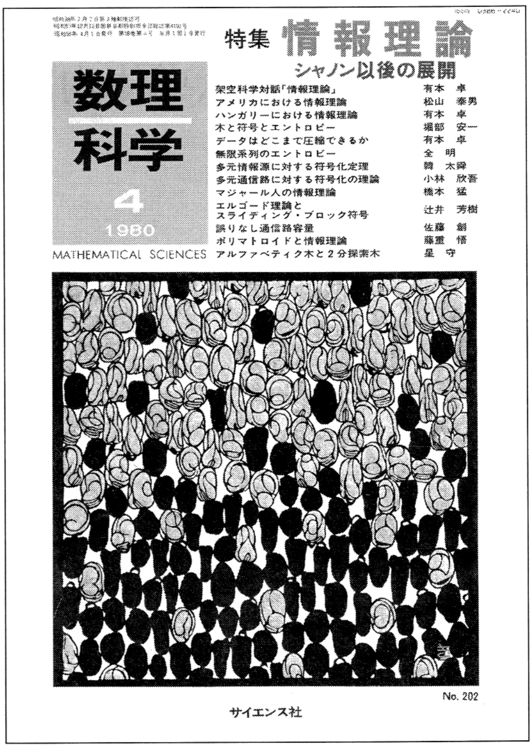
平井　武（京都大学名誉教授）

「数理科学」は語る

30年前から現代へのメッセージ

有本　卓

1980年4月号



情報理論の特集号が30年前に出た頃，筆者のヒーローはC.E. シャノンとA.M. チューリングであった．二人は，まさしく，情報科学とコンピュータサイエンスの基礎を創った．振り返ると，筆者の研究方向は1980年頃から，シャノン流の情報理論からはずれ，むしろシャノンが趣味として親しんでいたチェスプレーイングロボットやジャグリングロボットの世界に向い，それらを現実に技術展開することに集中し始めていた．研究室は機械工学科に所属し，機械系のロボティクス研究を開拓する任務を負っていたからである．

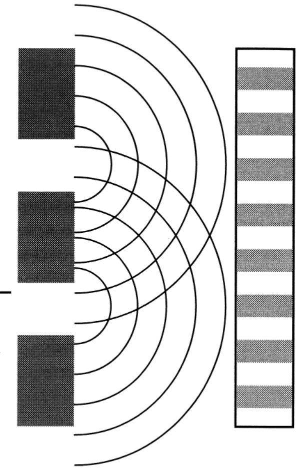
情報理論は，本特集号が出た1980年頃までは，シャノンが1948年に発表した長大な論文で提示した道筋を細密化することで発展したと言われる．振り返ると，シャノンが見通していなかった情報理論の新たな展開は，1980年以後になって，次々と生まれた．この新しい展開に筆者は何も貢献できなかったことに悔いは残るが，もう一つのロボティクス研究に責務があったことだけでなく，数理科学の新たな展開がこの未開の分野にあり得るという直観のほうに心を奪われていたからである．本特集号には，筆者のシャノンへの傾倒が架空科学対話に色濃く反映されているが，これはもはや昔の物語に過ぎなくなった．シャノン流を超えた情報理論の新たな展開はいつか新特集号として編まれるべきであろう．時機を逸しないように，色褪せることなく．

筆者の研究生活は，ほぼ半世紀（1959年に大学を卒業した）になるが，最後の10年間のヒーローはG.F.B. リーマンになっている．情報理論に熱中していた1970年代，リーマン幾何学やリーマン予想で有名なリーマンのことは，ほとんど知らなかった．ロボットの力学と制御の研究が始まったのは本特集号が編まれた1980年前後であるが，そのとき，ロボティクスに新たな数理科学的展開があり得るとは誰も思ってはいなかった．21世紀に入った頃，やっと，取り扱う対象物体の輪郭曲線や曲面の幾何学とロボットの手先を力学的に拘束させつつ，多自由度ロボットの姿勢全体の集合をリーマン多様体として見る方向が展開できるようになった．この新たなリーマン幾何学こそが，曲面とロボットの運動と制御を融合させ，Neo-Newtonian Mechanicsを造る基礎となるのではないだろうか．かくして，筆者は，今や，人の手やロボットハンドの巧みさ，器用さを数理的に解析する手段があり得ると思うようになった．リーマン幾何学は，ロボティクスによって数学的にもより“富み，かつ広汎”になるのではなかろうか．

（ありもと・すぐる，立命館大学総合理工学研究機構）

目で見て学ぶ量子力学

第17回　実験でひもとく量子の不思議



超伝導体内部の磁束量子を見る

外村　彰

17.1 はじめに

1989年，電子波を用いて磁束量子の姿を，初めて捉えることができた．とは言っても，超伝導内部の磁束量子の姿を見た訳ではなく，超伝導体外部に出てきた磁力線を見たのである．超伝導体内部の磁力線の観察が可能になれば，超伝導体内部の欠陥を同時に見ることができ，“磁束量子が欠陥と相互作用” する様子が目のあたりに見えるようになる．臨界電流[傍注1]の大きな実用材料の開発指針を得る上で手助けになる．

何より，超伝導体内部に閉じ込められた磁力線の様子をじかに見ることができるだけで，面白い．南部陽一郎先生が見抜いた “対称性の自然的破れ” も “クォークの閉じこめ” の原理も磁力線の形で観測できる．

これを可能にするには，まずは超伝導薄膜を透過する干渉性の良い高エネルギー電子線が必要になる．この夢は，思わぬチャンスから実現することになった．

[傍注1]　超伝導体は抵抗ゼロなので，無損失で電流を流すことができる．ただ，電流によって磁束量子が力を受け動き出してしまうと，超伝導状態が破壊されてしまう．磁束量子を動かさない工夫が必須になる．これが磁束ピン止めである．最大の超伝導電流のことを臨界電流と言う．

17.2 日立基礎研究所

1985年，日立は創業75周年を迎え，記念事業の一環として基礎研究所が設立されることになった．これは，当時の副社長の渡辺宏さんの10年以上にわたる念願であった．そして，35万ボルトホログラフィー電子顕微鏡の開発資金7億円は，研究の目玉として含めてもらった．

基礎研究所の組織はすぐにでき，未来を担う研究所にふさわしい土地探しから始まった．土地探しは難航した．新しい電子顕微鏡の設置条件として，地盤の固い浮遊磁場[傍注2]の少ない場所が選ばれた．日本中の候補地の中から，埼玉県鳩山の林野庁の試験場の跡地が選ばれた．建屋が15万坪という広大な土地

[傍注2]　電子線は磁場で容易に曲げられてしまう．とくに 10^{-8} ラジアンもの平行度の良い電子線にとって，周囲に浮遊している交流磁場は最大の敵である．高圧電線や蛍光灯の下などでは，大きな交流磁場が存在する．



図 17.1　日立基礎研究所（埼玉県鳩山町）.

に（図 17.1）完成したのは 5 年後の 1990 年だった.

35 万ボルトホログラフィー電子顕微鏡も 1985 年に開発がスタートし，建屋ができる前年の夏に完成，独立の電子顕微鏡棟の中に設置されて稼動し始めた.

この装置で，まず高分解能電子線ホログラフィーに取り組むことにした．その実験が完了した直後から，磁束量子の観察実験に着手し，1992 年，超伝導体サンプル中に一列に並んだパチンコ玉らしきものが一斉に動く姿を捉えた．初めは，その気にならないと見えないくらいのかすかなものだったが，実験条件を探っていくうちに，みるみるきれいな像になっていた．この像は，まさしくニオブの膜を突き抜ける磁束量子の像であることが分った．事実，この像は，臨界温度を越えると，とたんに姿が消えてしまう．我々は，“初めて見る磁束量子の動く姿” を息を呑んで見つめた．この後の発展は，あっという間だった．磁束量子が，まさにパチンコ玉のように見えるようになった.

渡辺さんは，磁束量子を見たかった張本人であり[傍注3]，私が電子線ホログラフィーの研究をスタートするきっかけを作ってくれた恩人でもある．そもそも私が日立の電子顕微鏡グループに入ったのも，渡辺さんに憧れてのことである．この連載記事の構想を練っていた 2007 年 7 月 17 日，渡辺さんは他界された．思えば，この 40 年間私は渡辺さんの手のひらの上にあったような気がする.

[傍注 3]　私が日立に入社した 1965 年頃，渡辺さんは磁束量子の観察実験を行っていた．第 15 回を参照.

[傍注 4]　鉄，コバルトなどの磁石になる材料でできた薄膜のことを “磁性薄膜” と呼ぶ．磁気テープや磁気ディスクなどに応用されている.

[傍注 5]　磁性薄膜は，膜全体が 1 つの磁石になっているとは限らない．いくつもの異なった向きの磁石に分かれていることがある．この 1 つ 1 つの小さな磁石のことを “磁区” と呼び，磁区の分布を “磁区構造” と呼ぶ．外部から大きな磁場をかけると，磁化の向きが一方向に揃って一様な磁石になる．磁場を変えたときに磁区構造が変化する様子は実用材料にとって重要な情報である.

17.3 ローレンツ顕微鏡法

超伝導薄膜内の磁束量子は，当初，磁力線の形で見る計画だった．「それが，なぜパチンコ玉になってしまったのか？」

実は，実験を始めたものの，それらしきものがさっぱり見つからない．そこで，まずは “ローレンツ顕微鏡法” で，磁束量子の存在を見つけようということになった．ローレンツ顕微鏡法は，磁性薄膜[傍注4]の磁区構造[傍注5]の観察用に使われてきた手法である．磁束量子に比べれば格段に大きな物体である．磁

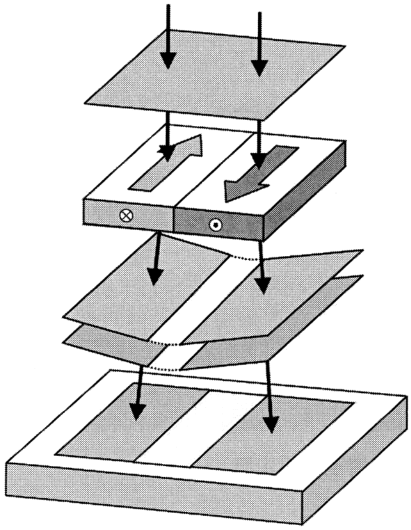


図 17.2　ローレンツ顕微鏡法の観察原理.

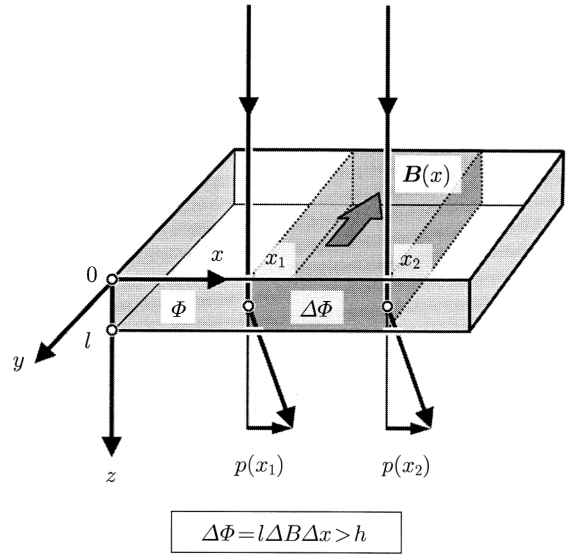


図 17.3　磁性薄膜を通過した電子の位置 (x) と運動量 (p).

束量子は，電子の位相変化にして，わずか $\frac{1}{2}$ 波長以下．磁束量子は，不確定性原理で決まるローレンツ顕微鏡法のちょうど観察限界であることも示されていた．「こんな小さなものをローレンツ顕微鏡法で果たして観察できるか？」—そこから話を始めることにする．

(a) 古典的説明

逆向きの磁区を持つ強磁性薄膜に電子線を入射させる（図 17.2）．電子線は，薄膜内部の磁場によってローレンツ力[傍注6]を受けて，少しだけ方向を変えて出ていく．1つの磁区を透過した電子線は，偏向された下方の面で，一方向にずれる．逆向きの2つの磁区を通った電子は，反対方向にずれるため，磁区と磁区の境界である“磁壁”に沿って，電子線が全くやって来ない領域ができる．つまり，図 17.2 に示したように，磁壁は暗い縞として観察できる．2つの磁区の磁場が逆向きの場合，今度は電子線が重なり合って明るい縞になる．こうして，明暗の縞として磁壁が観察できる．これがローレンツ顕微鏡法の観察原理である[傍注7]．

[傍注6] 電荷をもった粒子が磁場 $(\boldsymbol{B})$ の中を通るときに加わる力は，“ローレンツ力”と呼ばれている．力の大きさは磁場 $(\boldsymbol{B})$ と粒子の速度 $(\boldsymbol{v})$ に比例し，その向きは粒子の速度と磁場のいずれにも直角になる．数学的に表現すると，$-e(\boldsymbol{v}\times\boldsymbol{B})$ で与えられる．

[傍注7] 像のピントが外れているので，像をはっきりと見ようとしてピントを合わせてはならない．磁気コントラストは消失してしまう．

(b) 観察限界

「この方法で，どこまで微細な磁気的構造が観察できるのだろうか？」—1966年，D. ヴォレーベンは，この問題を議論し[1]，つまるところ，電子の位置と運動量がどれだけ正確に観測できるかという“不確定性原理”に帰着することを明らかにしている．この議論は，面白いので，ここで紹介しておこう．

薄膜内部の磁場 $\boldsymbol{B}(x)$ は，x だけの関数で y 軸方向を向いているとする（図 17.3）．厚さ方向には磁場分布は同じだとすると，速度 $\boldsymbol{v}$ で入射した電子は薄膜中で $-e(\boldsymbol{v}\times\boldsymbol{B})$ なるローレンツ力を受ける．薄膜の厚さ l が小さいと，入射電子は，ごくわずか向きしか偏向されない．z 軸方向に入射した電子線が

x 方向の運動量 $p(x)$ をもらって出ていくとすると，

$$\frac{\mathrm{d}p(x)}{\mathrm{d}t} = -evB(x) \tag{17.1}$$

となるので，$p(x)$ は，

$$\begin{aligned} p(x) &= -e\int_0^{t_0} vB(x)\mathrm{d}t = -e\int_0^{t_0} \frac{\mathrm{d}z}{\mathrm{d}t}B(x)\mathrm{d}t \\ &= -e\int_0^{l} B(x)\mathrm{d}z = -eB(x)l \end{aligned} \tag{17.2}$$

となる．$B(x)$ に比例した角度だけ偏向され電子線は下方の観察面で横にずれ，ローレンツ顕微鏡像として観察されることになる．

この方法では，磁場分布が一様で，$B(x)$ が x によらない場合には，すべての電子が一様に横ずれするだけなので，何も見えない．磁場が空間的に変化していたとき，どこまで正確に磁場分布 $B(x)$ が観察できるかを議論する．

2 点 x_1, x_2 に入射した電子が，$p(x_1)$, $p(x_2)$ の運動量をもらって，それぞれ向きを変えて出ていくとする．z だけ下方の面での強度分布が異なってきて，$B(x)$ の変化を検出できる．

不確定性原理によれば，電子の運動量 $p(x)$ と座標 x を同時に精度良く決めることはできない．つまり

$$\varDelta x \varDelta p > h \tag{17.3}$$

である．$p(x_1), p(x_2)$ の差 $\varDelta p = p(x_1) - p(x_2)$ を認識できる限界は，この不確定性原理で与えられることになる．$\varDelta x = x_1 - x_2$, $\varDelta p = p_1 - p_2$, $\varDelta B = B(x_1) - B(x_2)$ とおくと，(17.2) 式から $\varDelta p$ は $\varDelta p = -e\{B(x_1) - B(x_2)\}l = -e\varDelta Bl$ となる．ここで $\varDelta B = B(x_1) - B(x_2)$ である．$x = 0$ から $x = x$ に至る領域の磁束を $\varPhi(x)$ とすると 2 点 x_1, x_2 での磁束[傍注8]の変化量 $\varDelta\varPhi$ は $l\varDelta x\varDelta B$ で与えられるので $\varDelta x\varDelta p$ は，$\varDelta x\varDelta p = -\varDelta x\varDelta Bel = -e\varDelta\varPhi$ で与えられる．

こうして 2 点 x_1, x_2 での不確定性原理 (17.3) 式は，

$$|\varDelta\varPhi| > \frac{h}{e} \tag{17.4}$$

となる．つまり，$\varDelta x$ で指定される領域を貫ぬく磁束量 $\varDelta\varPhi$ をローレンツ顕微鏡法で測定しようとしても，原理的限界によって，$\frac{h}{e}$ 以下の情報は得られないことになる．この限界は，入射粒子の電子の速度によらない基本的な原理的限界である．「果たしてローレンツ顕微鏡法で $\frac{h}{2e}$ の磁束量子を観察できるのだろうか？」—正しくは，シュレディンガー方程式を解かなければならない．実際にシミュレーションを行ってみると，磁束量子 1 個の観察は可能であることが分かる[傍注9]．

そこで，次にシュレディンガー方程式を用いて得られた結果を紹介しておこう．

(c) 量子力学的説明

電子が $\boldsymbol{B}(x, y, z)$ なる磁場分布をもつ試料を透過したときの電子の波動関数

[傍注 8]　ある面積（A）を考えて，この面積を通り抜ける磁場（$\boldsymbol{B}$）の面積分が磁束になる．$\boldsymbol{B}$ の垂直成分 B_{A} と A をかけ合わせた値が磁束になる．

[傍注 9]　ただし，これは磁束量子が入射電子線に垂直に横たわっている場合についての考察である．膜を垂直に貫く磁束量子を見たい場合には，電子線に力が働かないので，膜を傾けなければならない．

$\Psi = \exp\left(\frac{iS}{\hbar}\right)$ を求める．ただし，電子線が磁場によって受ける効果は十分に小さいとする（WKB 法）．試料を通過して，z だけ後方の面での位相分布 $\frac{S(x,y,z)}{\hbar}$ は，次式で与えられる[2)]．

$$\frac{S}{\hbar} = \frac{e}{\hbar}\int_{-\infty}^{z} A_z(x,y,z')\mathrm{d}z' - \frac{ie}{2p}\int_{-\infty}^{z}(z-z')(\mathrm{rot}\,\boldsymbol{B})_z\mathrm{d}z'. \qquad (17.5)$$

A_z はベクトルポテンシャルの z 成分．第 1 項は位相変化を表し，第 2 項は，虚数 i がついていることから分かるように，位相ではなく，強度変化を与える．つまり，薄膜を透過直後の電子の波は，(17.5) 式の第 1 項で示されている位相変化が生じるが，後方の面で観察するローレンツ顕微鏡像には強度変化が生じる．強度分布には磁場分布が反映されている．

$\boldsymbol{B}$ の分布が空間的に一様で，変化がなければ $\mathrm{rot}\,\boldsymbol{B} = 0$ なので，第 2 項はゼロになって，強度変化は生じない．それだけではない．$\boldsymbol{B}$ の“渦度”がゼロ（$\mathrm{rot}\,\boldsymbol{B} = 0$）の場合にも，何のコントラストも生じない．真空中の磁場も，渦なしで調和関数的なので，$\mathrm{rot}\,\boldsymbol{B} = 0$ となり，コントラストが生じない．磁壁近傍のように，$\mathrm{rot}\,\boldsymbol{B}$ がゼロでない値を持つ場所で，コントラストが生じることになる．干渉顕微鏡像として，第 1 項による磁束量子の位相分布[傍注10]を高精度で測定すれば，磁束量子の投影された A_z 分布が観測できることになる．

[傍注10] 勿論，観測できるのは，位相そのものではなく，2 本の電子軌道に沿っての位相差である．磁束量子の磁束は $\frac{h}{2e}$ なので，$\frac{1}{2}$ 波長の位相変化を与える．

17.4 磁束量子はパチンコ玉

いよいよローレンツ顕微鏡法で磁束量子を観察する．超伝導薄膜に磁場をかけると，通常，磁束量子は膜に垂直に生じる．磁場と電子線が平行だと，相互作用しないので，薄膜試料を傾けなければならない（図 17.4 左）．薄膜を角度 θ だけ傾けた場合の位相変化は，2θ で与えられ，$\frac{1}{2}$ 波長以下になる[傍注11]．

[傍注11] 理由は，次の通りである．膜の表面からもれる磁束は磁束量子の磁束と逆向きになり，打ち消すことになる．計算してみると，膜を傾けた角度を θ とすると，位相差は 2θ で与えられる．例えば膜を $\frac{\pi}{4}$ (45°) だけ傾けたときには磁束量子の位相差は $\frac{\pi}{2}$，つまり $\frac{1}{4}$ 波長になる．

磁束量子は位相物体なので，ピントのあった電子顕微鏡像では観察できない．

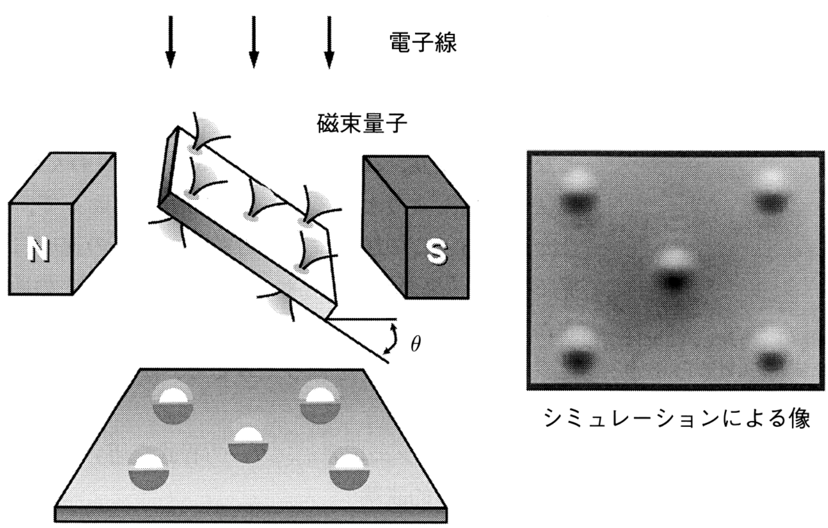


図 17.4　磁束量子のローレンツ顕微鏡像．

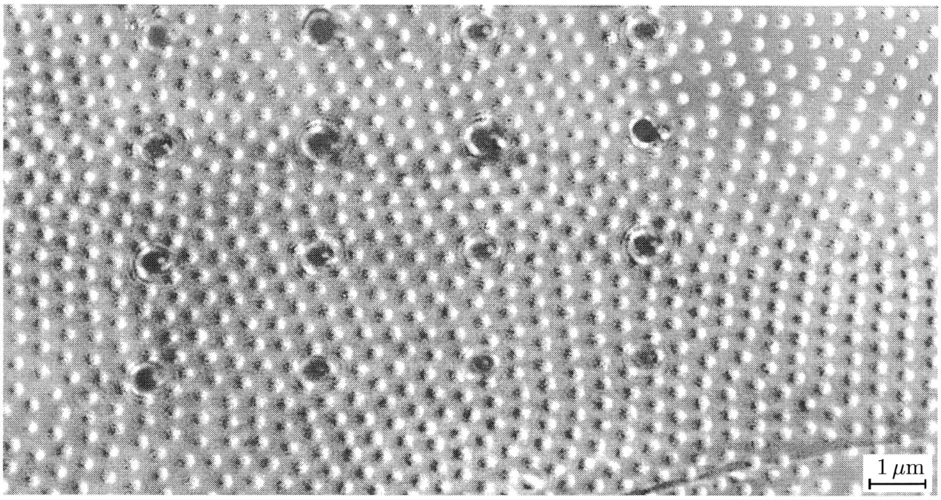


図 17.5　磁束量子（白黒のスポット）が点欠陥（黒いスポット）に捕捉されたため，きれいな三角格子に整列できない．

ピントをずらすと，位相変化が強度変化に変換されて，白黒のスポットとして観察できる．白黒の境界線の方向は磁束量子の磁場の方向に一致する．図 17.4 右にシミュレーション像を示してある．

これまで，静止している姿ならビッター法などで観察できた．しかし，磁束量子の動いている様子は，誰も見たことがなかった．ローレンツ顕微鏡法の場合，直接観察できるため，リアルタイムで捉えることができる．実際，この方法で観察した磁束量子の動きは，あたかも“生き物”であるかのように面白い振舞いをする．

しかも，この方法では，像はボケてはいるものの，欠陥を同時に見ることができる．このため，磁束量子がピンニングセンター[傍注12]に捕捉されたり脱離する様子を，動的に観察できるので，磁束ピン止め現象を直接見ることができる．

人工的な欠陥を導入すると，磁束量子は，さらに面白い振舞いをする．欠陥が点在しているニオブ薄膜中での磁束量子の写真を図 17.5 に示す．黒いスポットが，細く絞った Ga イオン[傍注13]ビームを薄膜に照射して作った点欠陥である．パチンコ玉の様に見える白いスポットは，磁束量子である．

磁束量子はお互いに反撥し合うが，磁場がかかっているので，たくさんの磁束量子がギュッと押し込められて三角格子状に並びたがる．しかし写真を見ると，磁束量子は点欠陥の位置に優先的に捕捉されるので，規則正しく格子状に並びたがっているが，点欠陥の存在でそうもいかない様子が見てとれる．

［傍注 12］ 超伝導体の中に欠陥などがあると，そこは超伝導性が弱かったり，常伝導だったりするので，そこに磁束量子が落ちこんで，動きにくくなる．これを“ピンニングセンター”と呼ぶ．

［傍注 13］ Ga のイオン源は容易にできるので，イオンビームを細く絞ってニオブ膜にあてると，その部分のニオブが削られて，小さな欠陥を作ることができる．

17.5 マッチング効果を見る

もっと密に周期的な点欠陥を作ると，さらに面白い．ある領域に四角格子状に欠陥を作ると，そこでは磁束量子は四角格子を組む（図 17.6）．図中の小さな丸印が点欠陥の位置を示す．欠陥のない領域では，磁束量子は三角格子とい

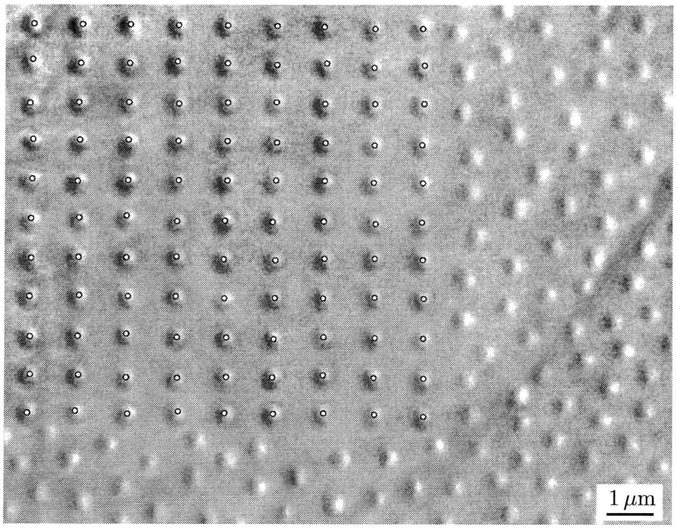


図 17.6　高密度の人工点欠陥が部分的に存在するときの磁束量子 ($T = 5\,\mathrm{K}$).

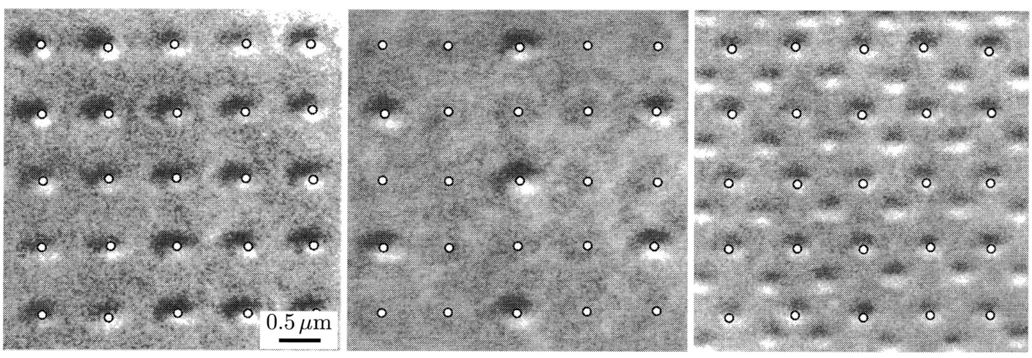


図 17.7　四角格子を組む人工点欠陥によって格子状に並ぶ磁束量子 (a) $H = H_1$, (b) $H = \frac{1}{4}H_1$, (c) $H = 2\,H_1$.

うよりランダムに点在している．ニオブの膜に弱い欠陥がたくさん存在し，ピンニングセンターとして働くためである．温度を T_c に近づけると，このピン止め効果は小さくなり，三角格子を組むようになる．このとき，特定の磁場で臨界電流が大きくなる現象（マッチング効果）が生じる．

「どうして，臨界電流が増すのだろうか？」—そのミクロなメカニズムを調べる目的で行った実験結果[3] が図 17.7 である．

欠陥と同じ密度で磁束量子が生じるマッチング磁場 H_1 をかけたとき，磁束量子は欠陥に 1 個ずつ入り，きれいな四角格子を組む（図 17.7 (a)）．このとき，臨界電流が大きな値をもつことは容易に理解できる．1 つの磁束量子が熱エネルギーをもらって欠陥から脱離しようとしても，移動すべき安定な空席は元の席以外にない．どうしても動かそうとするなら，もっと大きな力を加えて，たくさんの磁束量子を同時に動かす必要がある．

同じ現象は，マッチング磁場の整数倍やその分数倍のときにも起こる．例え

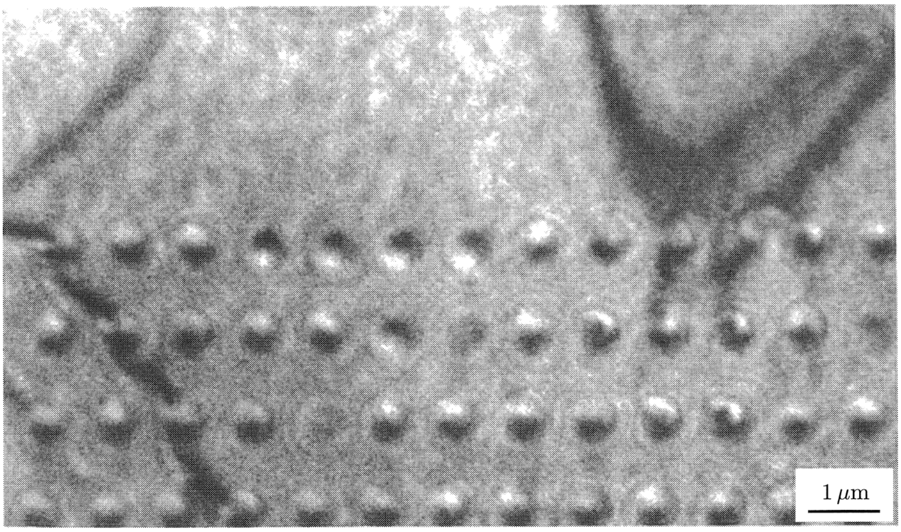


図 17.8　人工欠陥にトラップされた磁束量子と反磁束量子.

ば $\frac{1}{4}$ の磁場では，磁束量子は，水平方向 4 つおきの欠陥に入って規則正しい四角格子を組む（図 17.7 (b)）.

「マッチング磁場よりも大きくなったときには，どうなるだろうか？」—今度は 2 倍の磁場のときには，磁束量子はすべての“すきま”に入って，四角格子を組む（図 17.7 (c)）.

このようにして，マッチング磁場，その整数倍，そしてその分数倍のときに，磁束量子が規則正しく並んで固い格子を組み，臨界電流を大きくするミクロなメカニズムが直接観察できるようになった.

図 17.6 の状態でかけていた磁場を切ると面白い現象が起こる．点欠陥にピン止めされている磁束量子には変化がないが，それ以外の領域の磁束量子は膜の外に出てしまう，それだけではない．膜の縁から逆向きの磁束量子が入ってきて欠陥にトラップされる．上向きの磁束量子に下向きの磁束量子が衝突すると，対消滅をして，空孔になる．そこにもう 1 つの下向きの磁束量子がやってくると，下向きの磁束量子がトラップされる．その途中の経過を示したのが図 17.8 である.

17.6 次の狙いは高温超伝導体！

我々は，次に高温超伝導体の観察を狙った．動作温度が高かったり，層状構造のために，ピン止めが効果的でないため，当初期待していたよりも実用化が進んでいなかった．ところが，この磁束量子の観察を試みても，なかなか観察できない．高温超伝導体の磁束量子は，ニオブよりも一桁も太い．このため，磁場は 2 桁弱くなる．さらに膜は磁束量子の太さよりも厚くしなければならないが，35 万ボルトの電子線では，そんな厚い膜を透過するこができない.

一旦あきらめた“100 万ボルトホログラフィー電子顕微鏡への思い”が再び

浮上してきた．しかし，バブルがはじけてしまった後なので，予算獲得からして大仕事であった．

しかし，この装置は高温超伝導体を観察するためだけの装置ではない．超伝導以外にも，ミクロの世界の極限的な観察・計測手段として，様々な可能性を開いてくれるものと期待できる．$\frac{1}{100}$ 波長以上の精度で位相情報を検出することによって，ナノ領域に至る磁場や電場が観測可能であり，量子力学の基礎現象も観察できるようになる．短かい波長を利用してこれまでにない空間分解能でナノの世界を直接観察することもできる．ひょっとしたら原子モデルを見ているかのように原子を見たり，特殊な形をした原子の波動関数の様子を立体的に見たりできる可能性も生まれる．

巨大な望遠鏡を使って観察した遠い宇宙の美しい画像が我々の夢をかき立ててくれるのと同様，極めつけの電子顕微鏡によって，もっと身近なミクロの世界を高い分解能で見られるようになる．そうなればナノサイエンス，ナノテクノロジーの分野で有力なツールとしてさらに広く利用できるようになる．今，収差補正レンズと電子線ホログラフィーを使った，100 万ボルトホログラフィー電子顕微鏡を超える，極めつけの超高分解能電子顕微鏡の開発を計画している所である．2009 年度の補正予算で提案し，一旦は最先端プロジェクトとして認可されたが，仕分けによって予算が $\frac{1}{3}$ になってしまったので，今，開発できるか否かの瀬戸際に立たされている．この記事が印刷される頃には，どちらかに決まっているはずである．

参考文献

1) D. Wohlleben: *Phys. Lett.* **22** (1966) 564–566.
2) A. Fukuhara, K. Shinagawa, A. Tonomura and H. Fujiwara: *Phys. Rev. B* **27** No.3 (1983) 1839–1843.
3) K. Harada, O. Kamimura, H. Kasai, T. Mastuda, and A. Tonomura: *Science* **274** (1996) 1167–1169.

（とのむら・あきら，(株)日立製作所基礎研究所／(独)理化学研究所基幹研究所／(独)沖縄科学技術研究基盤整備機構）

リレー連載

現象から方程式を創り出す

第3回 —立式で鍛える論理的思考法—

振り子の運動

振動の方程式を創る

佐藤 光

1. はじめに

ピサ大学の医学生であった若き日のガリレイは，ある日の夕暮れ時，散歩の道すがら町の大聖堂に入っていった．薄暗い堂内では明かりが灯されたばかりらしく，天井からつり下げられたシャンデリアが大きく右に左に揺れていた．それを見たガリレイはとっさに自分の脈をとって，揺れの周期を測ってみた．**振り子の等時性**はこうして発見されたと一説には伝えられている．

ひもやバネにつるされた物体の規則正しい振動は，身の回りでよく見られる現象であるため，物理の学習では早い段階で必ず取り上げられる．また，固体中ではその構成要素である原子や分子がバネ振動と同じような振動をしている．比熱などの物質の諸性質は原子・分子のこのような運動を量子力学的に取り扱うことによって理解される．さらに，場の量子論においては電子や光子などの粒子は量子化された振り子の**量子**として表される．このように,「振り子の運動」はニュートン力学から量子力学，場の量子論に至る物理のAからZまでを通して奏でられる通奏低音にもたとえられる．

ここでは，バネや振り子の運動，減衰振動などを取り上げ，主にニュートン力学の範囲でこれらを理解するための方程式をいかに組み立てるかを考察する．このことを通して，様々な振動現象に通底する基礎方程式の存在を確認し，振動現象の統一的理解を目指す．

2. 運動の法則

物理現象を見て，それから方程式を創り出すためには，まず注目する現象の根底にある物理法則は何かを明確にとらえる必要がある．力学現象においては，それは**ニュートンの運動法則**である．特に，運動を記述する方程式を立てるための基礎になるのが，いわゆる**運動の第2法則**である．

運動の第2法則　物体（質点）の運動状態が変化するとき，その加速度 $\boldsymbol{a}$ は質点にはたらく外力 $\boldsymbol{F}$ に比例し，質点の質量 m に反比例する．

比例係数を K とすれば，この法則は

$$\boldsymbol{a} = K\frac{\boldsymbol{F}}{m} \tag{1}$$

と表される．ここで質点とは，大きさが無視できて，その属性として質量のみを考えればよいような物体のことである．また，太文字 $\boldsymbol{a}$ などはベクトル量を表すものとする．

長さ，質量，時間の単位として，それぞれ m, kg, s を用い，質量 1 kg の質点に $1\,\mathrm{m{\cdot}s^{-2}}$ の加速度を与える力を **1 N**（ニュートン）と定義するこ

とにより，(1) の比例係数は $K=1$ となる．こうして運動を記述する運動方程式は

$$m\boldsymbol{a}=\boldsymbol{F} \tag{2}$$

によって与えられることとなる．

3. 単振動

3.1 バネ振動

図 1 のように，長さ l のバネに質量 m のおもりを付けてぶら下げると，バネは少し伸びてつり合う．この伸びを Δl とすれば，バネが元に戻ろうとする力は伸びがあまり大きくない限り Δl に比例し，その大きさは鉛直上向きに $k\Delta l$ である．これを**フックの法則**と言い，比例係数 k は**バネ定数**と呼ばれる．おもりには重力 mg（g は重力加速度）が鉛直下向きにはたらくから，つり合いの位置での力の関係は

$$k\Delta l=mg \tag{3}$$

である．ただし，バネの質量はおもりに比べて十分小さく，無視できるものとする．(3) よりバネの伸びは $\Delta l=mg/k$ と決まる．

次に，バネをつり合いの位置から少し伸ばして手を放すと，おもりは上下に振動する．つり合いの位置を原点にして鉛直方向に，下向きを正として x 軸を設定する．バネの伸びが x のときのおもりにはたらく力の総和 F は，(3) の関係を用いると

$$F=mg-k(\Delta l+x)=-kx \tag{4}$$

となる．このように，おもりにはたらく力はつり合いの位置からのバネの伸び x（$x<0$ のときは縮み）だけで決まる．おもりの運動の加速度 a は

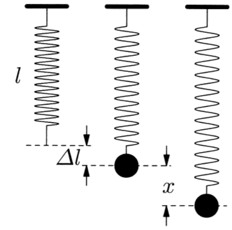


図 1　バネにつるされたおもりの振動．

$$a=\frac{d^2x}{dt^2} \tag{5}$$

であるから，バネ振動の運動方程式は (2) から

$$m\frac{d^2x}{dt^2}=-kx \tag{6}$$

と求められる．

おもりは質量が大きいほど，その運動状態を維持しようとする．このような性質を**慣性**と呼ぶ．運動方程式 (6) の左辺はおもりの慣性による力を表しており，**慣性力**と言う．また，(6) の右辺はバネが元に長さに戻ろうとする**復元力**である．この式は，振動現象が慣性力と復元力のバランスによって生じることを示している．(6) の形の微分方程式で表される運動を**単振動**，または**調和振動**と言う．

さて，長さや時間など，物理で扱われる量は数値に適当な単位を付して表される．このような量を**物理量**と言う．物理量の単位を [物理量] という記号で表すことにする．例えば力 F の単位は

$$[F]=\mathrm{N}=\mathrm{kg}\cdot\mathrm{m}\cdot\mathrm{s}^{-2} \tag{7}$$

である．バネ定数の単位は (4) より

$$[k]=\frac{[F]}{[x]}=\mathrm{kg}\cdot\mathrm{s}^{-2} \tag{8}$$

となる．そこで

$$\omega_s=\sqrt{\frac{k}{m}} \tag{9}$$

という量を考えれば，$[\omega_s]=\mathrm{s}^{-1}$ である．u を単位を持たない単なる数，すなわち無次元量とすれば，

$$t=\omega_s^{-1}u \tag{10}$$

と表され，ω_s^{-1} はバネ振動の時間のスケールを決める量であることがわかる．

同様にして，x_0 を振動の最大振幅とすれば，$[x_0]=\mathrm{m}$ であり，おもりのつり合いの位置からのずれ x は無次元量 f と次の関係にある．

$$x=x_0f. \tag{11}$$

(10), (11) の関係を用いて運動方程式 (6) を書き直せば，

$$\frac{d^2f(u)}{du^2}+f(u)=0 \tag{12}$$

が得られる．

方程式 (12) には単位を持った物理量は一切現れず，純粋な数学の方程式である．この形の微分方程式は未知関数 $f(u)$ に関して 1 次式であり，**2 階線形微分方程式**と呼ばれるタイプのものである．(12) の解は三角関数によって次のように与えられる．

$$f(u)=A\sin(u+\varphi). \tag{13}$$

この式で，A と φ は**積分定数**と呼ばれる任意定数である．再び (10), (11) を用いて物理量の間の関係に戻せば，(13) より

$$x(t)=x_0\sin(\omega_s t+\varphi) \tag{14}$$

が得られる．ここでは，x_0 を最大振幅としたので，任意定数 A は $A=1$ である．もう一つの任意定数 φ は時間の原点をどこにとるかによって決まる．x_0,φ は運動の初期条件によって決まる量である．また，(14) の形から $\omega_s/2\pi$ は振動数という意味を持つ量であることがわかる．

3.2 単振り子

図 2 のように長さ l の糸の先に質量 m のおもりを付けたものを**単振り子**と言う．糸を鉛直線からずらしておもりを放すと，おもりはつり合いの位置 O (最下点) のまわりで，半径 l の円周上を左右に揺れる．おもりには重力がはたらく他に，糸の張力 T がはたらく．張力はおもりが落下しないように支えているが，おもりの運動方向に対して

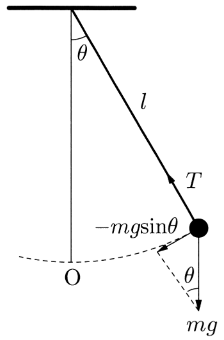


図 2　単振り子．

常に垂直にはたらくので，おもりの運動には影響しない．

このようなおもりの運動を記述するために，つり合いの位置 O を原点として，円周上に座標を設定する．振れの角度が θ のときのおもりの位置座標は原点 O からの円周上の距離 $x=l\theta$ である．なお，振れの角度 θ は O を通る鉛直線を基軸にして反時計回りを正として測ることとする．

図 2 に示すように，おもりには鉛直下向きに重力 mg がはたらく．おもりをつり合いの位置に引き戻そうとする力 F は重力の接線方向の成分であるから

$$F=-mg\sin\theta \tag{15}$$

である．また，円周上を運動するおもりの加速度は

$$a=\frac{d^2x}{dt^2}=l\frac{d^2\theta}{dt^2} \tag{16}$$

で与えられる．したがって，単振り子の運動方程式は (2) により

$$ml\frac{d^2\theta}{dt^2}=-mg\sin\theta \tag{17}$$

となる．

この方程式は未知関数 $\theta(t)$ に関して 1 次ではない．このようなタイプの方程式は**非線形微分方程式**と呼ばれる．しかし振れの角度 θ が小さい場合には，$\sin\theta$ を

$$\sin\theta\approx\theta \tag{18}$$

と近似することにより，(17) は次のようにバネ振動で得られた方程式 (6) と全く同じ形の**線形微分方程式**になる．

$$l\frac{d^2\theta}{dt^2}=-g\theta. \tag{19}$$

これから先の議論はバネ振動の場合と同様である．

$$\omega_p=\sqrt{\frac{g}{l}} \tag{20}$$

とすれば，ω_p の持つ単位は $[\omega_p]=\mathrm{s}^{-1}$ であり，ω_p^{-1} は単振り子の運動の時間スケールを決める．そこで θ_0 を振れの最大角度として

$$t=\omega_p^{-1}u,\quad \theta=\theta_0 f \tag{21}$$

とおき，無次元の変数 u および f を用いて (19) を書き直せば，(12) と同じ微分方程式が得られる．

3.3 電気振動

ここでは力学を離れて，図 3 のような電気回路を考えよう．コンデンサーとコイル（ソレノイド）を直列につないだ回路は **LC 回路**と呼ばれ，以下に述べるような興味深い振舞いを示す．

初めにコンデンサーに電荷を与え，スイッチを閉じると電荷はコイルを通じて放電される．回路を流れる電流は放電が進むにつれて減少するが，その際，電磁誘導によってコイルの両端には電流の変化を妨げるように起電力が発生する．そのために電流は流れ続け，コンデンサーは以前と逆向きに帯電するに至る．すると今度は，コンデンサーから逆向きに電荷が移動し始める…．こうしてコンデンサーの極板の間には交互に正負の電圧が発生する．回路を流れる電流も交互にその向きを変えて流れ，あたかも振り子に見られるような振動現象が発生する．この現象を**電気振動**と言う．

電気容量 C のコンデンサーに電気量 Q が帯電しているときの極板間の電位差は Q/C である．一方，電流 I が変化するときにインダクタンス L のコイルに発生する起電力は $-LdI/dt$ であるから，図 3 の閉回路について次式を得る．

$$-L\frac{dI}{dt}=\frac{Q}{C}. \tag{22}$$

回路を流れる電流 I と電気量 Q の間には $I=dQ/dt$ の関係が成り立つ．これを (22) に用いることにより

$$L\frac{d^2Q}{dt^2}=-\frac{1}{C}Q \tag{23}$$

を得る．これは，単振動の方程式 (6) や (19) と同じ形をしているから，この電気回路においてもバネや単振り子と同じ振動現象が見られるのである．

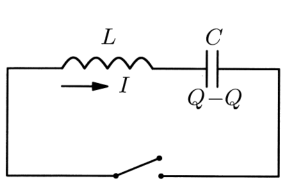


図 3　LC 回路．

コイルに発生する起電力は電流の変化を妨げ，流れを維持しようとするから，慣性力に対応している．コンデンサーは帯電した電荷を放電することによって元の電気的に中性の状態に戻ろうとするので，極板間の電位差は復元力に対応する．すなわち (22) は慣性力と復元力のバランスを表していて，それゆえに電気的な振動が生じることとなる．

この電気振動の振動数は $\omega_e/2\pi$,

$$\omega_e=\frac{1}{\sqrt{LC}} \tag{24}$$

によって与えられ，ω_e^{-1} は電気振動の時間スケールを決めていることもバネ振動などの場合と同様である．また，Q_0 を回路に最初に与えた電気量とし，$Q=Q_0f, t=\omega_e^{-1}u$ とおいて，無次元の変数 u, f を用いて (23) を書き直すことにより，(12) と同じ方程式が得られることも容易に確かめられる．

4. 減衰振動

空気中で振り子を振らすと，次第に振幅が小さくなり，やがては止まってしまう．これは空気などの流体中を物体が運動するときには，抵抗力がはたらくからである．この現象を**減衰振動**と言う．抵抗力の大きさは，物体の速さが小さいときはほぼ速さに比例することが知られている．そこでバネ振動について，この抵抗力の効果を考えよう．

バネに付けたおもりにはたらく力は，バネの復元力 (4) の他に，抵抗力

$$F_r=-rv=-r\frac{dx}{dt}\ (r>0) \tag{25}$$

がおもりの運動方向とは逆向きにはたらく．したがって，抵抗力を考慮した運動方程式は (6) の代わりに次式で与えられる．

$$m\frac{d^2x}{dt^2}=-kx-r\frac{dx}{dt}. \tag{26}$$

この方程式を無次元の変数で書き直すために，(9)～(11) の他に，

$$r=2m\omega_s\gamma \tag{27}$$

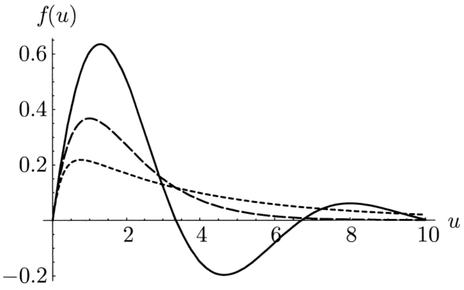


図 4　微分方程式 (29) の解の様子（$\gamma > 1$ 点線，$\gamma = 1$ 破線，$\gamma < 1$ 実線，初期条件：$f(0) = 0, f'(0) = 1$）．

によって無次元の定数 γ を導入する．実際に，r の持つ単位は (25) の関係により，

$$[r] = \frac{[F_r]}{[v]} = \mathrm{kg} \cdot \mathrm{s}^{-1} = [m][\omega_s] \qquad (28)$$

であるから，(27) で定義された γ は無次元である．こうして (26) は次の微分方程式に帰着する．

$$\frac{d^2 f(u)}{du^2} + 2\gamma \frac{df(u)}{du} + f(u) = 0. \qquad (29)$$

ここで $\gamma = 0$ とおくと，抵抗力のない場合の単振動の方程式 (12) が再び得られる．

微分方程式 (29) は定数係数の 2 階線形微分方程式であり，その一般的な解法が知られている．図 4 に γ の大きさを変えた場合の解の様子を示す．この方程式はバネ振動だけでなく，抵抗力のはたらく単振り子や電気振動の場合でも，それぞれの運動方程式を無次元の変数で表すことにより共通して成り立つ**振動現象の基礎方程式**である．電気振動の場合の抵抗力は電気抵抗によって生じるので，図 3 の回路に抵抗を直列に加えた **LCR 回路**を考えればよい．

5. ポテンシャルエネルギーと力

この節では振動の方程式を少し別の視点から導いてみよう．

簡単のために 1 次元運動を考え，その**ポテンシャルエネルギー**を $U(x)$ とする．安定なつり合いの位置を $x = 0$ とすれば $U(x)$ はそこで極小となるから，$U'(0) = 0, U''(0) > 0$ である．したがって $U(x)$ を $x = 0$ のまわりでテイラー展開すれば

$$U(x) = U(0) + \frac{1}{2}U''(0)x^2 + \cdots \qquad (30)$$

となる．エネルギーの原点を $U(0) = 0$ となるようにとれば，安定なつり合い点の近傍では

$$U(x) = \frac{1}{2}U''(0)x^2 \qquad (31)$$

と近似できる．

一方，質点に力 F がはたらいて微小な変位 $\varDelta x$ が生じたとき，力のした仕事は $F\varDelta x$ である．このとき，**エネルギー保存則**により，ポテンシャルエネルギーの変化 $\varDelta U$ は

$$\varDelta U(x) = -F\varDelta x \qquad (32)$$

となる．$\Delta x \to 0$ の極限を考えることにより，力 F はポテンシャルエネルギー U と

$$F(x) = -\frac{dU(x)}{dx} \qquad (33)$$

によって関係づけられることがわかる．力が 3 次元ベクトル $\boldsymbol{F}$ の場合には (33) はポテンシャルエネルギー U の**勾配**で与えられる．

$$\boldsymbol{F} = -\nabla U. \qquad (34)$$

安定なつり合い付近の運動では，ポテンシャルエネルギーは (31) で与えられるから，質点にはたらく力は (33) より

$$F = -U''(0)x \qquad (35)$$

である．こうして力が決まったので，運動方程式は (2) から次のように求められる．

$$m\frac{d^2x}{dt^2} = -U''(0)x. \qquad (36)$$

これは第 3 節で学んだ単振動の方程式に他ならない．

以上の考察は一般の次元，あるいはより複雑な力学系にも拡張することができるが，ここでは**二重振り子**の例を取り上げる．

図 5 に示した振り子は，長さ l の糸に質量 m のおもりをつるし，それにさらに長さ l の糸で質量 m のおもりをつるしたものである．この振り子が鉛直面内で微小振動するときの運動方程式を導く．

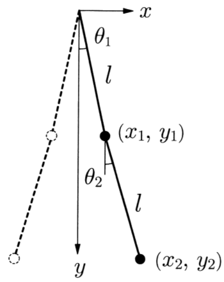


図 5　二重振り子.

振れの角度 θ_1, θ_2 は小さいので，三角関数の近似式 (18) および

$$\cos\theta \approx 1 - \frac{1}{2}\theta^2 \tag{37}$$

を用いる．この近似で，おもりの x 軸方向の変位は

$$x_1 = l\theta_1, \quad x_2 = l(\theta_1 + \theta_2) \tag{38}$$

である．y 軸方向の変位は (37) により x 軸方向の変位に比べて高位の微小量なので無視する．

ポテンシャルエネルギー U は (37) の近似をして

$$\begin{aligned} U &= mgl\left(\theta_1^2 + \frac{\theta_2^2}{2}\right) \\ &= \frac{mg}{l}\left\{x_1^2 + \frac{1}{2}(x_2 - x_1)^2\right\} \end{aligned} \tag{39}$$

と表される．ここで，(38) の関係を用いて次のような変数変換を行った．

$$\theta_1 = x_1/l, \quad \theta_2 = (x_2 - x_1)/l. \tag{40}$$

質点にはたらく力はポテンシャルエネルギー (39) に (34) を適用して得られるが，多変数の場合には微分が**偏微分**になる．こうして運動方程式は

$$\begin{aligned} m\frac{d^2x_1}{dt^2} &= -\frac{\partial U}{\partial x_1} \\ &= -m\omega_0^2(3x_1 - x_2), \end{aligned} \tag{41}$$

$$\begin{aligned} m\frac{d^2x_2}{dt^2} &= -\frac{\partial U}{\partial x_2} \\ &= -m\omega_0^2(x_2 - x_1) \end{aligned} \tag{42}$$

によって与えられ，これらは次の**連立微分方程式**で表される．ただし，$\omega_0^2 = g/l$ である．

$$\frac{d^2}{dt^2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\omega_0^2 \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{43}$$

ここではこの方程式の解法には触れずに，結果のみを記す（A, B, α, β は積分定数）．

$$\begin{aligned} x_1(t) = A\sin(&\omega_+ t + \alpha) \\ &+ B\sin(\omega_- t + \beta), \end{aligned} \tag{44}$$

$$\begin{aligned} x_2(t) = (1 - \sqrt{2})&A\sin(\omega_+ t + \alpha) \\ &+ (1 + \sqrt{2})B\sin(\omega_- t + \beta). \end{aligned} \tag{45}$$

$\omega_\pm$ は**固有角振動数**と呼ばれ，次式で与えられる．

$$\omega_\pm^2 = (2 \pm \sqrt{2})\omega_0^2. \tag{46}$$

こうして，二重振り子の運動は単振動の組合せになっていることがわかった．

力は大きさと方向を持ったベクトル量であるが，ポテンシャルエネルギーは大きさのみを持つスカラー量である．したがって複雑な系ではまずポテンシャルエネルギーを計算し，それから (33) または (34) によって力を求める方法が有効である．

以上の考察から，次の注目すべき結論が得られる．

> 安定なつり合い付近での運動は，常に単振動か，あるいは単振動を組合せたものである．

単振動（調和振動）が物理の様々な場面に現れる主な理由はここにある．

6. さらに勉強するために

この小論では振り子の運動に代表される単振動について，現象のメカニズムの考察に基づいて運動方程式をいかに立式するかを中心に取り上げた．得られた方程式の解法や解についての解析は割愛したので，これらについては参考文献などによる読者の発展的な学習を期待したい．

参考文献

1) 佐藤　光：『新版 物理数学ノート』（サイエンス社）．
2) 佐藤　光：『微分方程式の基礎と応用』（サイエンス社）．
3) 稲見武夫：『常微分方程式』（岩波書店）．
4) 薩摩順吉：『物理の数学』（岩波書店）．
5) 和達三樹：『物理のための数学』（岩波書店）．
6) 長岡洋介：『振動と波』（裳華房）．

（さとう・ひかる，兵庫教育大学名誉教授）



数学の道しるべ
～研究者の道とは何か～
第15回

確率論との出会い

楠岡 成雄

もし，あなたの研究分野は何かと問われれば「stochastic（確率的）なものすべて」と答えるつもりでいる．研究者になるまでのこと，そして研究者になってからどのようなきっかけでそのときの研究テーマを選ぶことになったかについて，思い出すままに書いていく．

1. 小中高時代

小学生の頃，算数は大変得意だった．特に，パズルのような問題は面白がってよく解いていた．高学年になったとき，数学の面白さを伝えてくれる本に出会った．本のタイトルはうろ覚えであるが『ピタゴラスから電子計算機まで』であったように思う．今でもはっきり覚えているのは，1～10までと2, 4, 8, 16, . . ., 1024までの数字が左右に並べて書いてあった表で，4と16の積を考えるとき，その左にある数字2と4の和は6で，6のところの右を見ると64とありそれが答えとなっている．すなわち，かけ算が足し算で計算できるという説明で，面倒なかけ算が計算しやすい足し算に変わってしまうということに非常に感動した．その後に計算尺について説明してあり，すべての数の積が扱えると書いてあった．不思議なものがあるものだと思ったことを今も覚えている．

中学生のときに『不思議な数学』（ノースロップ著）に出会った．この本では$1=0$を証明したり，すべての角度はゼロであることを証明したりしていて，注意深く計算を行わなかったり，図に頼るととんでもない結論に導かれることを教えてくれた．数学においては，何よりも議論の精密さ・厳密さが求められることを知り，曖昧な理解を排除するということが大事であると知った．ますます数学が好きになり，その後様々な本を読んだが，私にとって最も影響のあった本だと思う．高校生のときには大学では数学を勉強しようと決めていた．一方，中学生の頃から大学紛争が起き，私が中学3年のときに東大入試が中止された．我が校でも制服廃止運動が起き，色々な議論が起きたが結局制服が廃止された．高校生のときは1970年安保闘争の影響が我が校にも及び，様々な討論が交わされた．その結果，3日間授業がなくなり終日ホームルームが行われたこともあった．私は議論好きだったので積極的に討論などに参加した．同級生に誘われマルクスの読書会に参加したこともあった．ただ，私は生来のあまのじゃくなので次第に反発し，安保賛成を唱えて，近代経済学の入門書を買って一人で勉強したりもした．当時は社会に役に立つような職業に就くべきであるという風潮が強くあったが，近代経済学を少し勉強した結果，数学を通じて社会貢献はできると感じて

いた．今，思うと，中学高校で起きたことは，ただ社会で起きたことに共鳴していただけで，自分自身は本当に深く考えて行動してはいなかったように思う．当時，こんなごたごたは受験勉強の妨げになるだけだと冷めたことを言っている人もいた．しかし，この経験で社会に目を向けるようになり，視野が広がったように思う．

2. 大学時代

1972 年に東京大学に入った．1 年生の頃は麻雀ばかりしていたような気がする．その頃，1 年生向けの位相幾何学のセミナーがあり，同じ学年で数学科を目指しているかなり多くの人はそのセミナーに出席していたと後から聞いたのであるが，私は工学部のコンピュータのセミナーに出席していてそのような数学のセミナーがあることに気がつかなかった．当時，工学部に進学すると FORTRAN が必修なのであるが，私自身はプログラム言語を勉強することはこのセミナー以後はなかったので，今から思うとそのセミナーをとってよかったと思っている．もっとも，当時のプログラムは一行をカード 1 枚に機械でパンチして打たねばならず，短いプログラムでもカードが数十枚になってしまう．教育用計算機は 1 台しかなく，プログラムを通すのは順番待ちで係にカードを渡して早くても 10 分くらいは待たねばならない．私は，カードのパンチをしょっちゅう打ち間違えてプログラムエラーばかり出て，パンチ機械と計算機係の人との間を何往復もし，エラーなしで通すのに苦戦したことをよく覚えている．現在の PC は夢のような機械だと思う．

2 年生の冬学期になり希望通り数学科に進んだ．ここで初めて，数学科の同期と一緒に授業を受けることになった．数学科では演習の時間が印象的であった．出された問題を家で一生懸命解くのであるが，なかなか解けない．後で知ったのであるが，当時は助手の間に代々，演習問題が引き継がれており，その中には奇問難問も少なくなかった．どのくらい解けるかで，その年の学生の力がわかるということだったらしい．自分が解けなかった問題を他の人が前に出て黒板で解くと，素直にこの人はよくこんな解法を思いついたなと感心し，その人に一目置くようになった．我々の学年は自主セミナーが非常に盛んで，そのような人たちと一緒にセミナーをやるようになった．テキストはミルナーの写像度に関する講義ノートのような幾何学のものが多かった．特に，I 君と仲が良くなり，彼とは修士修了までに二人で自主ゼミを行うことが多かった．彼の選んだ河田敬義著：『代数曲線論』とか熊ノ郷準著：『擬微分作用素』とかを一緒に読んだ．その知識も後の研究に役立ったので，彼には感謝している．もっとも I 君自身は修士修了のとき，「怠惰な生活は送りたくない」と言って数学者にはならず厚生省の官僚になってしまった．

4 年生になるときにゼミを選ばなければならなかった．応用的なことを勉強したかったので，第 1 志望は偏微分方程式，第 2 志望は確率論を選んでアンケートを提出した．ところが確率論を第 1 志望にした学生は一人しかおらず，担当教官の伊藤清三先生は学生が一人ではゼミはやらないとおっしゃった．第 2 志望にしていたのは私一人だけで，ゼミを決める集まりでは，藤原大輔先生が「自分も確率論を勉強したかった」などとおっしゃり，その雰囲気に負けて確率論をやることになってしまった．テキストは伊藤清先生のオルフス大学での講義録であった．普通のコピーは存在したが高価だったので，青焼きのコピーを作ったことを覚えている．このテキストはとても明快で面白く，確率論を研究してみようかと思い始めた．ちょうどその頃，伊藤清先生が集中講義に来られ，確率解析の入門的な講義をされた．この講義で伊藤先生は細かいことは何も話されなかった．ただ $dB_t^2 = dt$ の説明を一生懸命されたことと，「確率論は無限次元空間の理論である」とおっしゃたことだけをよく覚えている．大学 3，4 年のときは経済学も勉強してみようと思い，経済学部の講義にも出てみたのであるが，数学科の講義が過密で講義は結局一つしかとれなかった．ただ，夕方に開講される経済学部の統計学の鈴木雪夫先生のゼミにはほぼ皆勤で参加した．このときに知り合った人たちとは

今でも毎年 1 回以上顔を合わせている．特に，1 年先輩だった国友直人さん（現東京大学大学院経済学研究科教授）には，卒業後も時々学問に関して教わる機会があった．私がファイナンスの研究を始めるきっかけを作ってくれた人でもある．現在，数学科の講義は駒場キャンパスで行われ，本郷キャンパスにある経済学部とは遠く離れている．そのため，当時のように数学科と経済学部の講義をとることは困難である．IT が問題を解決してくれるかもしれないが，総合大学はやはりすべての学部が同じキャンパス内にあるべきだと思う．

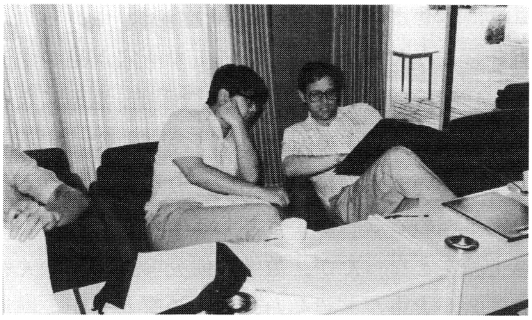
3. 大学院時代

1976 年に大学院に入った．専門は確率論を選び，上野正先生が指導教官となった．ゼミは駒場で行われ，毎週 1 回駒場に通った．当時，駒場には上野先生のほかに丸山儀四郎先生，高橋陽一郎さんが教官としておられ，院生も D1 の舟木直久さんを筆頭に私を含め 4 人いた．そこで，全体のセミナーを同じ曜日に毎週やろうということになった．その頃，丸山先生はブラウン運動の非線形汎関数の性質に強い関心を持たれ，諸性質，特に中心極限定理に関して精力的に研究をされていた．後になって考えてみると，この頃はマリアバン解析の現れる直前であり，丸山先生は全く違った角度から同じ対象を研究されていたこととなる．丸山先生は自身の極限定理について大学院の講義もされたが，毎回，膨大な計算をされ，黒板が一本の式で埋まるほどであった．まちまちの大きさの広告紙の裏に大量な式を書いて束ねたものを，先生が講義のためのノートとして使っておられたことを今でも覚えている．大学院 2 年のときに，今はチューリッヒ工科大学教授であるクンシュさんがスイスから丸山先生の所に 1 年間留学してきた．彼はマルコフ場の研究をしており，ゼミでもその話をしてくれた．彼の研究していたのはガウス系の場合で基本的に共分散関数の性質を調べて確率場のマルコフ性を示すのであるが，私には特別な場合はマルコフ性は明らかでガウス系である必要もないと思い，その話を彼にした．彼はそんなに簡単に示せるはずはないと言うので，きっちり証明をしようとしたら，証明に 1 カ所穴が見つかった．しかし，彼にできると言った手前，意地になって，あの手この手で穴を埋めようとした．するとナイトという人の結果が見つかりそれを使うとその穴を埋めることができてしまった．その内容が修士論文となったが，数学の研究ではとにかく粘ることが大事だと感じた．修士の終わり頃，同期である京都大学の重川一郎君から修士論文が送られてきた．その論文には驚くべきことが書かれていて，夢中になって読んだ．それは，当時まだプレプリントであったマリアバンの論文を，新しい概念を導入してわかりやすく整理し直したものであるが，自分が抱えていたある問題の解決方法も与えていた．マリアバン解析はその後，様々な人たちの手により急激に巨大な理論かつ道具へと成長していき，私の専門分野ともなるが，それが始まる瞬間に研究者となったことは大変に幸運であった．

4. 研究者となって

1978 年に数学科の助手となり，無限次元解析としての確率論という何だか曖昧なものを研究テーマとして掲げていた．その年，岡部靖憲先生が名古屋大学から数学科の助教授として移ってこられた．偶然にも岡部先生はマルコフ場の研究の第一人者であられた．岡部先生は粘り強いタフな先生で，私の研究について毎週話を聞いて下さりアドバイスを下さった．私は岡部先生を助手として手伝わなければいけないはずだったのだが，学生のゼミにはつきあわなくてよいと言って下さった．当時の数学科の助手は，年度内に半年だけ演習を 1 コマ担当すればよかったので，本当に研究に集中することができた．現在，若い人のパーマネントなポジションが減っており，一番研究ができる時期を不安定なポジションで過ごさねばならない．このことが若い人たちの研究を萎縮させていると思う．何とかせねばならないのであるが．

その頃，ドイツに行っておられた大阪大学の福島正俊先生から岡部先生経由でアルベベリオ–ホエフクローンの無限次元空間上のディリクレ形式



ストルックさんと（1982年）.

の論文のコピーを戴いた．この論文は量子場に関連した確率場を無限次元空間上の確率過程と解釈して，そこに福島先生が精力的に研究されたディリクレ形式の理論を適用しようとしたものである．物理の理論と結びついているために，あまり好き勝手な仮定がおけない．その一方で構造が深く，周辺には膨大な結果がある．是非，このようなことにチャレンジしたいと思い，まずはこの論文で不十分な部分を拡張して論文としてまとめた．一方，当時はまだマリアバン解析の研究を始めていなかったが，丸山先生の考えを進めて，ウィナー汎関数が解析的という特殊な場合を調べて，マリアバン解析ではわからないことについての結果が得られた．

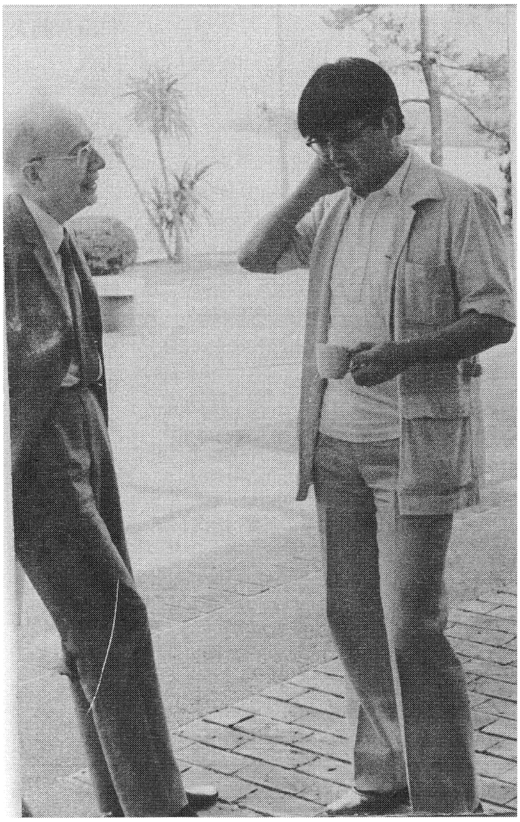
1982年に福島先生が中心となりマルコフ過程に関する谷口シンポジウムが開かれた．そこには数名の外国人研究者が招待されたが，招待者の中にドイツ・ルール大学のアルベベリオさん，アメリカ・コロラド大学のストルックさんがいた．私は幸運にもその研究集会に参加することができた．ストルックさんは「あなたの研究は面白いが方向が偏っている．今はまずマリアバン解析そのものを研究すべきだ．アメリカに招待するから一緒に研究しよう」と言ってくれた．しかし，突然なので，迷っていると，「わかった，伊藤と話をする」と言って，やはり研究集会に参加していた伊藤清先生のところに行ってしまわれた．翌日になり，伊藤先生から「アメリカに行きなさい」と言われて，アメリカに行かざるを得なくなってしまった．どうも，私の運命はいつも偉い人の一言に左右されてしまうようである．しかし，これがきっかけで1995年くらいまでストルックさんとマリアバン解析に関しての共同研究が続くことになった．彼とは年が10歳以上離れていたが何故か気が合い，短期間に結果が量産された．試しに調べてみたところ，共著論文は12編であるが，論文の延べページ数は495ページあった．

アルベベリオさんからもドイツに招待され，ドイツへは2度，計1年弱ほど行った．ドイツではデュールさんとも知り合いになり，統計物理学や量子場の理論について色々なことを学ぶことができた．そこで学んだ色々な問題に現在も取り組んではいるが，なかなかよい結果が出ていない．真の物理モデルにおいて数学的に厳密な結果を出すことは大変難しいことだと思う．ドイツで学習院大学の江沢洋先生とお会いし，帰国後は江沢先生のセミナーに3年ほど参加させて頂いた．しばらくして学習院大学の助手に服部哲弥君が着任し，最初のセミナーでフラクタルに関して講演をした．フラクタルという名前はもちろん知っていたが，具体的な問題に関する講演を聴いたのは初めてであった．非常に面白いと思い，フラクタル上の拡散過程に関する研究を始めた．

1987年に京都大学の数理解析研究所に移った．京都には木上淳君，熊谷隆君がいてフラクタル上の解析学に関して大きな業績が上がった．ただ，ケンブリッジ大学のバーロー氏を初めとする多数の外国の研究者も研究に参入し，結果が急速に積み上がると同時に対象がより複雑なものになりハードな解析が要求されるようになった．現在も熊谷君はハードな解析を使いこなし，業績を上げ続けているが，若い人は簡単に参入できないような状況が生まれているように思う．

5. 数理ファイナンスの研究

京都にいたときに経済学部のゼミの先輩である国友さんに論文に関する相談を受けた．そのとき，数理ファイナンスに関して色々と教えて頂いた．非常に発想が面白く，数学としてもかなり面白い対象ではないかと感じ研究を始めた．同じ頃，慶



マリアバン先生と（1982 年）.

応大学経済学部の丸山徹さんと知り合い，経済学者と数学者が一緒になって数理経済学の研究集会を毎年行おうということになった．研究集会を通して山崎昭さんを初めとする多くの経済学者と知り合いとなり，また研究集会で経済学に関する多くの講演を聴くことができた．ファイナンスに対する考え方もかなり変化していったように思う．残念ながら経済学の研究をするまでには至っていないが．

1993 年に東京大学に戻った．その頃，金融機関で働く中堅の人たちの中で本格的にファイナンスを勉強したいという人たちがいたので，彼らを博士課程の学生として受け入れた．数学や数理ファイナンスについて彼らに教えたが，セミナー後の食事の席ではむしろ彼らから，現実にファイナンスがどう用いられているか，どのような金融商品が開発されつつあるか，実際の金融機関の動向など様々なことを教わった．さらに，現在は東京大学大学院経済学研究科教授となっている高橋明彦君を金融機関より数理科学研究科の教官として招くことができた．社会人学生や高橋君のおかげで，実務をにらんだ数理ファイナンスの研究が可能となり，研究が偏った方向に進んで行かなかったと思っている（実際はどうなのかは自分ではわからないが）．

最近は特に計算ファイナンスという分野の研究を行っているが，かつて研究したマリアバン解析が重要な道具として使えることがわかってきた．ファイナンスの研究を始めた 1990 年頃はマリアバン解析が役に立つとは思われていなかったし，それを使って研究しようと考えたこともなかった．数学は，そして科学はどのような分野であってもどこかでつながっているのかもしれない（今ではマリアバン先生ご自身もファイナンスの研究を行っており，本まで出版している）．

思い出話に終始したが，自分自身の研究の歩みを考えると不思議なことに若い頃に勉強した，数学の研究とは関係がないと思われることが，ずっと後になって研究につながってきたように思う．最近の若い人たちを見ると（と言い始めると自分も年をとったとも感じるのだが）効率を考えすぎて無駄な勉強は一切しないという人が増えているように見える．しかし，どんな人でも 40 年，50 年先を見通せるはずもない．面白いと思うことは，数学に限らずどのようなことでも勉強することが重要で，その勉強が長い目で見れば数学の研究にも役に立つのだと思う．

（くすおか・しげお，東京大学大学院数理科学研究科）

特　集（2010年5月号予告）

物理学事始め

各分野へのガイド

本体952円

大学に入っての物理学	和田　純夫
ニュートンは何を見たか	江沢　洋
電磁気学	近重　悠一
流れを見る物理	生井澤　寛
熱と物性	田中　文彦
統計物理って何？	吉岡大二郎
量子力学	和田　純夫
相対性理論	二間瀬敏史
自然界の究極構造	日置　善郎

特　集（2010年3月号）

物理数学の効用

物理の具体例から入る数学抽象世界

本体952円

巻頭言	佐々木　隆
変分原理と物理法則	窪田　高弘
ベクトル解析と場 ベクトル，微分形式，位相不変量（電磁気学，場の理論）	綿村　哲
フーリエ解析と流体の波や流れ	船越　満明
巨大数コイン投げ問題 具体例で確率論の概念を体感し統計力学の基礎へと向かう	佐々　真一
ラプラス変換と揺らぎ	香取　眞理
線形代数と量子力学	筒井　泉
対称性　量子力学と群・代数の表現論	小竹　悟
グリーン関数と波動方程式 量子力学，場の理論とともに	静谷　謙一
時空の幾何学　相対性原理を極めよう	加藤　晃史
可積分系の数学　ソリトンを源流として	鈴木　淳史
《現象から方程式を創り出す》　第2回	大貫　義郎
《数学の道しるべ》　第14回	小谷　元子

編集後記

4年にわたって続けてきましたシリーズ特集も今回で最後となりました．小誌が2006年に日本数学会の出版賞を受賞したことを契機に数学啓蒙に役立つ特集を組もう，ということで進めてまいりました．如何でしたでしょうか．2007年4月号の［代数編］，2008年4月号の［解析編］，昨年4月号の［幾何編］，そして今回の［確率編］と全部で4つの特集があります．今回初めて手に取る読者の方がいらっしゃいましたら，過去，3つのものも是非，手に取ってご覧下さい．

この4年の間に映画や小説の中に数学が出てきたりして数学も身近になってきたかな，と個人的には感じております．数学の理論が使われているところは，皆さんの身の周りにもっと多くあると思います．是非，探してみて下さい．

ところで「数学が使われている」ということで，もう一つ，数学的な思考も大変重要なテーマであると思います．昨今は数学離れが深刻になっていると言われていますが，実は，数学で使われている論理的思考能力というのは，あらゆる場面で有用で，物事を考えるのにとても役に立ちます．最近では政治や経済の世界にも理科系出身の方が出てきたことがクローズアップされてきて，今後はこのあたりのところにも論理的な思考方法が浸透するのではないかと期待しております．(R)

数　理　科　学　4月号

2010年4月1日発行

数理科学編集部　　発行人　木　下　敏　孝
TEL. (03) 5474–8816
FAX. (03) 5474–8817

ホームページ　http://www.saiensu.co.jp
ご意見・ご要望は　sk@saiensu.co.jp　まで．

発行所　ⓒ　株式会社 **サ イ エ ン ス 社**　TEL. (03) 5474–8500（代表）
〒151–0051 東京都渋谷区千駄ヶ谷1–3–25　振替 00170–7–2387　(03) 5474–8600（広告部）

印刷・製本　三美印刷株式会社　山岡景仁